CRANFIELD UNIVERSITY

Sean Turner

Challenges for implementing water resources planning frameworks based on stochastic modelling assessments: the case for change in England and Wales

SCHOOL OF APPLIED SCIENCES

CRANFIELD WATER SCIENCE INSTITUTE

EngD

2014

Supervisor: Professor Paul Jeffrey

August 2014

CRANFIELD UNIVERSITY

SEAN TURNER

Challenges for implementing water resources planning frameworks based on stochastic modelling assessments: the case for change in England and Wales

SCHOOL OF APPLIED SCIENCES

CRANFIELD WATER SCIENCE INSTITUTE

EngD

2014

Supervisor: Professor Paul Jeffrey

August 2014

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Engineering

ABSTRACT

This research examines the case for change in the regulated water resources planning process in England and Wales. The primary contribution to knowledge is delivered through the identification of practical, conceptual and institutional challenges associated with emerging planning methods based on stochastic modelling assessments. Four alternative modelling trials are executed and then compared to existing practice using real-world water resources systems. In-depth, structured interviews capture the views of a range of practitioners closely involved in the planning process. The study finds that the trialled approaches are technically feasible and can be executed using existing models and freely-available data. This finding counters the widespread view—exposed during interviews—that water companies are ill-equipped to conduct stochastic modelling assessments. However, some of the purported benefits of these frameworks failed to materialise in the case analyses. The study identifies arbitrary assumptions that threaten the transparency and rigour of the emerging methods. The practitioner interviews highlight widespread scepticism and perceived business risks associated with a shift away from deterministic planning. The thesis also delivers a number of methodological developments and is structured using a simple, novel matrix that characterises water availability assessment methods according to the way performance is measured and the way hydrological uncertainty is treated.

Sponsors: UK Engineering and Physical Sciences Research Council, United Utilities, Anglian Water, Severn Trent Water.

Funding code: EP/G037094/1

ACKNOWLEDGEMENTS

The author thanks: Paul Jeffrey, academic supervisor at Cranfield University, for comments, advice, guidance and support; Jonty Dobson and Chris Matthews, industrial supervisors at United Utilities, for flexibility and support; Richard Blackwell and Mark Smith, for insights, data, and the chance contribute to United Utilities' water resources planning agenda; Steve Whipp and Nathan Johnson, for initiating this research; Steve Kaye and Mark Jones, for co-sponsoring; David Marlow, for giving me the opportunity to work at CSIRO; Will Clark and Peter Edgley at Oxford Scientific Software, for providing a licence for Aquator; Tania Baker, and all other staff and students on the STREAM IDC; and Neil Upton, Steve Moncaster, Martin Berry, Stuart Sampson, Polly Chancellor, Ken McDonald, Glenis Pewsey, Paul Merchant, David McGrath, Tom Nichols, Judith Stunell, Steve Wade, Gaynor Kenyon, Dave Champness, Chris Harris, Chris Lambert, Rob Sage, Bruce Rhodes, Udaya Kularathna, Keith Beven, Tom McMahon, Murray Peel, Paul Roberts, Lu Zhang, Francis Chiew, Andreas Efstratiadis, Marie Ekström, KS Tan, Ben Baker, Shiroma Maheepla, Tirusew Asefa, Edoardo Borgomeo, Ian Holman and Jerry Knox, for insights and interesting discussions.

This project was financially supported by the UK Engineering and Physical Sciences Research Council, United Utilities PLC, Anglian Water and Severn Trent Water through the STREAM Industrial Doctorate Centre. Any opinions, findings and conclusions or recommendations expressed in this thesis are those of the author and do not reflect the views of any of the organisations mentioned above.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AWBM	Australian Water Balance Model
CMIP3	Coupled Model Intercomparison Project phase 3
DC	Demand Centre
DO	Deployable Output
DWRS	Demand-Weighted Risk Score
GCM	Global Climate Model
EA	Environment Agency
ES	Emergency Storage
MAF	Mean Annual Flow
MAP	Mean Annual Precipitation
MAT	Mean Annual Temperature
MRS	Multi-Reservoir System
Ofwat	Office of Water Services
PET	Potential Evapotranspiration
PPE	Perturbed Physics Ensemble
PR	Price Review
RDM	Robust Decision Making
RRL	Rainfall Runoff Library
RRS	Relative Risk Score
SAC	Special Area of Conservation
SOSI	Security of Supply Index
SRES	Special Report on Emissions Scenarios
SRS	Single-Reservoir System
SSSI	Site of Special Scientific Interest
TSS	Total System Storage
UKCP09	UK Climate Impacts Programme 2009
UKWIR	UK Water Industry Research
VBA	Visual Basic for Applications
WAFU	Water Available for Use
WRMP	Water Resources Management Plan
WRZ	Water Resource Zone

CHAPTER 1 INTRODUCTION

1.1 Background

Water service providers are responsible for the secure and efficient supply of water to households, businesses, public premises and industry. This requires maintenance and operation of upstream bulk supply systems—herein termed "water resources systems"—that abstract, store and transfer water in such a way that ensures continuity of supply through severe and prolonged droughts. These systems also need to be continually upgraded to cope with changing conditions, which may include: increasing demand for water driven by population growth and demographic change; changes to abstraction licence conditions and compensatory release requirements driven by environmental legislation; projected impacts of climate change on the natural availability of water; and changing expectations of customers and other stakeholders. Deciding how and when to upgrade a system in response to these developments is the discipline of water resources planning.

In England and Wales these responsibilities fall on the private water companies, which are obliged to "develop and maintain an efficient and economic system of water supply" (Water Industry Act of 1991). Their efforts in realising this goal are highly regulated: each company has a duty to prepare and maintain a Water Resources Management Plan (WRMP) that adheres to detailed planning guidance set out by the Environment Agency (Water Act of 2003). The current guidance prescribes a planning methodology that follows classic "capacity expansion" principles: forecast supply, forecast demand, buffer against possible error and uncertainty pertaining to both supply and demand using planning margins, and select the least cost combination of measures to balance supply and demand over a 25 year planning horizon (Environment Agency, 2012). The underlying metric of "supply", termed Deployable Output (DO), is normally determined computationally by simulating the water resources system under the historical recorded inflows (typically 50 – 100 year record). The simulations are repeated to iteratively converge on the maximum demand that the system can sustain without failure—which is triggered if storages are depleted to emergency levels or if desired levels of service are breached—and whilst satisfying abstraction licence conditions and environmental constraints.

The capacity expansion approach to water resources planning is relatively simple, uses readily available hydrological data and has served the economically-advanced countries well over the course of a century (Stakhiv, 2011). However, a number of academics have begun to question whether classic design principles can continue to serve the water resources planning community and broader society in future. A central concern expressed in the international literature is that the existing planning tools fail to adequately accommodate uncertainty—particularly uncertainty associated with anthropogenic climate change impacts on natural water availability (Milly *et al.*, 2008; Hirsch, 2011). The spectre of climate change diminishes the already-weak prospects for reasonable drought frequency estimation; there exists no model that can reliably predict how severe the droughts of the 21st century might be. This creates a dilemma for planners. The consequences of under-designing a water resources system range from undesirably frequent restrictions on customer water use (implemented through non-essential use bans) through to catastrophic system failure (i.e., running out of water), which carries substantial societal, environmental and economic risks. On the other hand, mitigating water shortage risks almost always implies developing expensive and controversial infrastructure projects—reservoirs, inter-basin transfer schemes, water re-use schemes, seawater desalination facilities, and so on. The water resources planner must somehow strike the balance between levels of investment and indeterminable water shortage risks, whilst simultaneously satisfying the many and diverse interest groups that might be affected by the decision.

In response to this challenge, a number of recent studies have developed and promoted analytical approaches that purport to help planners deal more effectively with hydrological uncertainties and other uncertainties, such as population growth and its impact on the demand for water. Invariably, these emerging planning frameworks rely on more computation: replace the historical recorded flows with stochastically-derived inflow sequences and then explore uncertainties through extensive simulation of the water resources system. The analysis exposes vulnerabilities in the system, which allegedly informs a more "robust" plan (Dessai *et al.*, 2009; Wilby *et al.*, 2010; Weaver *et al.*, 2013). This thesis explores the utility and practicality of these forms of analysis from the perspective of the water company. It seeks to identify the challenges that planners would face if they were to embrace planning frameworks that employ stochastic modelling assessments. In doing so, it provides a timely and important contribution to a burning industry debate, one that reflects a wider discussion taking place across the international water resources planning community.

1.2 Research contribution

The thesis contributes to the existing knowledge in three distinct ways. First, it develops and demonstrates a set of practical methods for executing stochastic modelling assessments with the software, models and data used routinely by water companies. Second, it develops a simple matrix for characterising water availability assessment methods, and then uses the framework to distinguish which types of analysis are most likely to yield new and valuable insights for planners. Third, it identifies a range of conceptual and institutional challenges associated with planning frameworks that employ stochastic modelling assessments. The research contests the notion that well-informed, transparent decisions are the natural outcome of the emerging methodologies.

1.3 Thesis layout

The thesis is delivered in “portfolio” format, comprising five distinct peer-reviewed journal-standard articles, presented in Chapters 3 – 7 and outlined in Table 1.1.

Table 1.1 Journal articles presented in the thesis

Article title [thesis chapter]	Journal	Status
Risk-based water resources planning in England and Wales: Challenges for execution and implementation [3]	Urban Water Journal	Published
Challenges for informing water resources planning decisions through 'robust decision frameworks' [4]	Water Resources Management	Under review
Linking climate projections to performance: an application of the decision scaling approach on a large urban water resources system [5]	Water Resources Research	Published
Standardizing traditional water availability assessments by correcting the reserve storage bias [6]	To be confirmed	Ready to submit
Industry views on water resources planning methods—prospects for change in England and Wales [7]	Water and Environment Journal	Accepted for publication

Chapter 2 identifies the gap in knowledge through a concise review of literature, which is kept short to avoid repeating details provided in the introductory sections of the separate articles. A simple research framework is introduced to provide an overarching structure to the separate studies and to help define the research scope and contribution. Research questions and methodology are presented.

Chapters 3, 4, 5 and 6 present four methodological approaches illustrated on respective case studies. Each is based on a modelling trial using a particular planning framework. **Chapter 7** describes an interview study that captured practitioner views on the strengths and weaknesses of existing planning methods and the case for change toward a stochastic approach. The separate articles are stand-alone studies and can be read in any order; they are presented here in the order that they were conducted.

Chapter 8 triangulates the insights from the five separate studies to directly address the research questions. **Chapter 9** concludes the thesis with a summary of findings, their significance, limitations and recommendations for future research.

1.3.1 Article presentation

The separate articles presented in Chapters 3 – 7 have been formatted to a consistent style for the purpose of this thesis (e.g., referencing format standardised, American spelling switched to English, minor grammatical adjustments, etc.), but are identical in content to the versions submitted to five international journals (Table 1.1). Co-author contributions are outlined in an introductory table that precedes each article in the thesis. Two papers have been accepted subject to revisions; the relevant thesis chapters incorporate these revisions based on the peer reviewers' recommendations. References are given in a single section at the end of the thesis rather than separately for each article.

1.4 Thesis constraints

The thesis is primarily concerned with the fundamental modelling assessment methods used to inform a set of emerging planning frameworks. The outputs include some methodological developments for executing these modelling assessments and then interpreting the results. However, the thesis is not concerned with developing a planning framework for decision making under uncertainty; instead it seeks to understand the utility and practicality of water resources planning frameworks that employ extensive (stochastic) vulnerability analysis.

Whilst three of the four case studies in the thesis employ models and data based on water resources systems in England and Wales, the theoretical contributions can be applied more widely. However, the thesis makes no assumption with regards to what processes and methods are already used in planning process outside of England and Wales. Therefore, any discussion on the "case for change" refers to the case for change in planning methods prescribed by the WRMP guidelines.

The thesis comprises five distinct papers that have been submitted to international journals. These journals demand that a paper can stand alone. As a result, the thesis may appear somewhat disjointed and perhaps repetitive through the middle section, with new introductory sections covering much of the same ground as in previous chapters. Also, detailed literature review and discussion for each analytical component of the work are delivered separately in the relevant papers. Chapter 2 therefore offers a chronologically structured thematic overview of analytic approaches to water resources planning as a context for the overall contribution of the thesis. The author recognises the impact the "portfolio format" has on the general flow of the writing. The discussion section in Chapter 8 has been designed specifically to amalgamate the relevant findings from each case study and bring together the overall contributions of the thesis.

The work described in this thesis was undertaken on the premises of an industrial sponsor, which directed the research toward new goals as findings emerged. As a result, the fourth article—presented in Chapter 6—may appear to veer significantly from the main objectives of the study. Specifically, this study did not trial a planning framework for decision making under uncertainty. However, the article complements the thesis in other ways. It demonstrates a stochastic analysis of one of the largest and most complex water resources systems in England and Wales, thereby lending to the conclusions around practicality of executing stochastic modelling approaches.

All of the case studies described in the thesis are based on conjunctive use water resources systems dominated by surface water resources, particularly reservoirs and river abstractions. The modelling methods developed and described may not be easily transferred to systems dominated by groundwater sources. The range of water resources system types available for study was constrained by the models available through the primary research sponsor.

1.5 Definitions of key terms

The thesis employs a number of technical terms that describe the performance of a water resources system. These terms are often associated with a variety of different meanings, so Table 1.2 sets basic definitions, which are loosely based on those proposed by Hashimoto *et al.* (1982a, 1982b). The means of deriving these metrics vary slightly throughout the study. More detailed definitions are supplied at the relevant stages of the thesis.

Table 1.2 Definitions of key technical terms of supply system performance

Reliability	The frequency with which a performance threshold is breached—could be based on a reservoir storage trigger or a level of demand shortfall.
Resilience	The time taken for a water resources system to recover from an unsatisfactory state, normally indicated by reservoir storage levels.

Vulnerability	The magnitude of demand shortfall experienced by customers under conditions of failure—can be expressed as ratio of water delivered to the target demand.
Risk	A measure of the performance of a supply system based on the probability and consequences of tangible hazards, such as supply restrictions.
Robustness	A measure of the adequacy of system performance under a wide range of scenarios as opposed to optimal performance under a single scenario; ability to cope with many alternative futures at minimal cost.

CHAPTER 2 LITERATURE REVIEW AND RESEARCH FRAMEWORK

2.1 Concise review of traditional and emerging approaches to water resources planning under uncertainty

Perhaps the first analytically rigorous approach to designing a water resources system was performed in England in the late 19th century (such an approach would have been near-impossible before then because planners would have lacked hydrological data). Wenzel Rippl (1883) plotted the cumulative streamflows entering a reservoir to develop what became known as the mass curve method—a graphical technique for determining the storage capacity required to sustain a given demand without failure under recorded conditions. This simple approach became the global standard for reservoir design. Varlet (1923) adjusted the technique such that it would determine the "ideal flow" for a given storage—a process akin to what today's engineers would call yield analysis.

The engineers of the early 20th century recognised the fallibility of their methods: the brevity of the available inflow records meant their designs would be vulnerable to relatively minor droughts. They attempted to remedy the problem by synthesising hydrological time series, first by (literally) shuffling the recorded flows (which were marked on a deck of playing cards) and later by developing sophisticated models that could generate statistically coherent replicate sequences (Hazen, 1914; Sudler, 1927; Barnes, 1954; Maass *et al*., 1962). The original idea was that streamflow time series could be characterised as a stationary stochastic process such that a large enough sample would contain a near fail-safe design drought. But the premise for this approach unravelled with the discovery of long-term hydrological persistence and nonstationarity (Hurst, 1951), which ultimately meant that planners would have to cope with the uncomfortable reality that the climate can shift abruptly and unpredictably; there exists no well-defined envelope of hydrological variability or quantifiable upper limit on the severity of drought that a given catchment or set of catchments may experience (Klemeš, 1987). Well-informed hydrologists and engineers would recommend designing water resources systems with conservative planning margins to provide additional robustness (Matalas and Fiering, 1977), although in England and Wales it took a severe drought in the mid-1990s to awaken the industry to the real prospect of more severe events than those experienced in the last century. The drought of 1995/96 was particularly severe in the north of England, where a combination of prolonged dry conditions and, in Yorkshire, slow management response led to reservoir failures that raised an impetus for more conservative design standards that were mandated in subsequent planning regulations (Uff, 1996; Department of Environment, 1996).

The history serves as a reminder that the fundamentals of planning have changed little in a century. Certainly water resources systems are more complex, and yield is determined computationally rather than graphically. But the basic design principles are the same (ensure yield exceeds demand) and the longstanding problem persists: how to design the system given the severe uncertainty of future climate conditions?

The majority of recent method development studies in this field begin with reference to Milly *et al.* (2008), who declared hydrological stationarity "dead" in the wake of anthropogenic climate change. The authors called on the hydrological community to revamp water resources analytic planning methods to embrace a new reality of severe uncertainty. Whilst a number of prominent hydrologists have disputed the idea that severe uncertainty is new (e.g., Lins and Cohn, 2011; Matalas, 2012), there can be no doubt that the climate change adaptation agenda has motivated the academic community to develop a range of methods for dealing with uncertainty in water resources system design. The question is whether these methods can help solve the problem at hand and traverse the gap between theory and practice.

The key tenets of the emerging planning methods are "robustness" and "flexibility." The goal is to design for satisfactory performance over a broad range of futures, which is purported to be a divergence from the tradition of aiming for optimal performance in a single predicted future (Dessai *et al.*, 2009; Lempert and Groves, 2010; Wilby, 2010; Gober, 2013; Weaver *et al.*, 2013). Given the aforementioned planning margins—which in England and Wales translate to "Emergency Storage" and "Headroom", as well as some other less obvious planning heuristics—the proposition that traditional practice has sought to optimise for a predicted future is debatable. "Humility," and "common sense" might be more appropriate descriptors of the traditional design ethos than "predict and provide" (Kundzewicz, 2011; Lins and Cohn, 2011). Of course, any planner could in theory design a system to cope with all manner of uncertain climate and demand scenarios; to do so would create a robust *system*, but would perhaps not constitute a robust *decision* given the substantial opportunity costs (financial and socio-environmental) associated with over-designed infrastructure. The challenge is to devise a robust investment plan that considers all sources of potential regret across the range of uncertain futures.

The literature contains dozens of analytical approaches to help achieve this aim in a water resources planning context. Recent suggestions that seek to nurture flexible, robust planning decisions include "decision scaling" (Brown and Baroang, 2011; Brown *et al.*, 2011; Hallegatte, 2012; Moody and Brown, 2013; Ghile *et al.*, 2014)), "Robust Decision Making" (Groves *et al.*, 2008; Lempert and Groves, 2010; Stakhiv, 2011; Chen *et al.*, 2013), "real options analysis" (Jeuland and Whittington, 2014), "risk-based planning" (Hall *et al.*, 2012a; Hall and Borgomeo, 2013) and various multi-objective optimisation approaches (e.g., Rosenberg, 2012; Kasprzyk *et al.*, 2013; Giuliani *et al.*, 2014).

At the heart of all these emerging methodologies lies an extensive vulnerability analysis of the water resources system, involving hundreds, or perhaps thousands, of simulations of the system under a wide range of plausible future hydrological conditions. This form of modelling assessment is not new, and neither is its application in water resources decision analysis, although its impact on real-world planning has been limited. Three decades after the birth of "systems analysis" (i.e., the use of optimisation and stochastically-varied flows in water resources system design), Rogers and Fiering (1986) found that practitioners preferred the more basic approach of "identifying several alternative plans... [and] subjecting them to all (or a critical part) of the historical hydrological record." The authors' prediction that "increased level of utilization will be seen... as the current availability of small computers exercises its full impact" has been invalidated by a further three decades of experience. "Systems analysis," and the use of stochastically-varied flows more generally, is absent from the bulk of water resources planning documents, not only in England and Wales, but across the developed world (e.g., Rush *et al.*, 2011; New Jersey Department of Environmental Protection, 2011; Tennessee Department of Environment and Conservation, 2013; Water Research Foundation, 2013).

Apparently the practitioner community has been reluctant to embrace more complex forms of analysis. Perhaps these methods are deemed too difficult or too demanding on staff time (Lund, 2008). Perhaps modelling studies have been overshadowed and rendered unimportant by non-technical concerns, such as the preferences of interest groups and their probable reactions to particular decisions (Loucks, 1992). Or maybe the practitioner community has been underwhelmed and unconvinced by the prospects for generating valuable insights through extended analysis. They would be in good company: Myron Fiering (pioneer of stochastic hydrology) found "systems analysis" to be almost worthless for solving problems of water resources planning under uncertainty (Fiering, 1976). Vit Klemeš delivered a similarly brutal synopsis: "...mathematical models for risk analysis, decisions under uncertainty, etc., have been further advanced, their theory refined, and their divergence from reality has often reached what seems to be a point of no return: they have become an end in themselves, intellectual parlour games played behind a façade of practical-looking jargon" (Klemeš, 2001).

The Klemeš critique might not immediately undermine the prospects for the new cadre of analytical methods listed above. Arguably, the aims of modelling and analysis have progressed in recent years. Emphasis has shifted toward achieving *transparency* under conditions of uncertainty, such that the emerging methods focus on exposing the nature of alternative designs by explicating assumptions, highlighting trade-offs and expanding the range of risk information available to the planner. A risk-based approach that clearly explicates the benefits, costs and trade-offs implied by different plans is said to have "normative appeal" (Hall and Borgomeo, 2013). Robust Decision

Making is said to “help decision makers understand the vulnerabilities of their plans” (Lempert and Groves, 2010). An info-gap approach is said to provide “a rich variety of information to support adaptive management” (Korteling *et al.*, 2011). A multi-objective robust decision making approach is said to help planners “characterize the most important vulnerabilities in their systems” (Kasprzyk *et al.*, 2013). The analysis need not recommend the decision—instead it helps the planner identify “promising solutions” with due consideration to the full weight of plausible outcomes (Rosenberg and Madani, 2014).

The recent wave of methodological developments has not gone unnoticed in England and Wales, where a selection of academics and consultants have advocated and promoted wider use of extensive vulnerability analysis through stochastic modelling (Hall *et al.*, 2012a; Dessai *et al.* 2013). The industry now appears to be in the early stages of a discussion on whether and how to implement these tools as part of the WRMP process (e.g., UKWIR 2014/15 Project WR02 – “Water Resources Management Plan 2019 Methods”). So far this discussion has suffered a lack of rigorous case demonstration targeted at identifying and characterising the practical and conceptual implementation challenges, particularly from the perspective of a water company (Arnell, 2011; UK Water Industry Research, 2012a; CH2MHill, 2013). This thesis aims to address this knowledge gap through four separate case studies, which are based on real-world water resources systems, and which adopt the modelling tools and data currently used by company analysts. The study seeks to develop some ways of deploying the emerging approaches and to identify the challenges for implementing them in a practical setting.

2.2 A simple matrix for characterising water resources modelling methods

Given the many and diverse planning approaches being developed, tested, discussed and promoted, the task of thoroughly investigating their practicality and utility seems near-impossible. How could a particular approach or set of approaches be selected from the many in order to begin the case analyses? Conveniently, all water resources planning methods—old and new—share fundamental traits that should allow for some important questions to be tackled without having to investigate the nuances of each individual approach. The goal of this section is to characterise these fundamental traits in a simple matrix that can then be used to define the scope of the research and amalgamate the separate articles that form the thesis.

2.2.1 Defining planning frameworks by modelling assessment procedure

The practice of water resources planning comprises a number of separate tasks—collection of data, customer research, project appraisal, communication with stakeholders, regulatory reporting, etc. At the heart of the process lies the modelling assessment, which influences most other component parts. It determines what data will be required, what project appraisal methods will be available, how the state of the system will be interpreted and communicated and, ultimately, how planning decisions will be made and justified. The data derived from the modelling assessment are propagated through the planning process, as depicted in Figure 2-1. Understanding the nature of the water resources modelling assessment is a pre-requisite to understanding any particular planning approach—and all prospective planning approaches are built around such assessments.

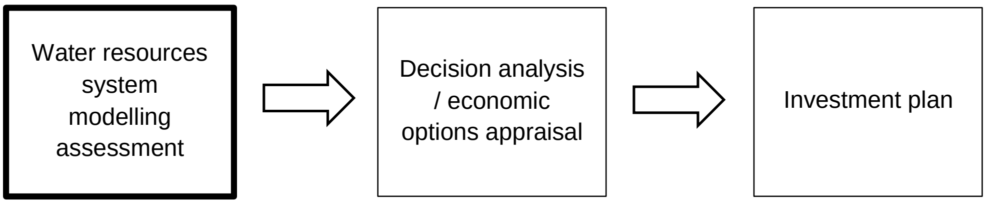


Figure 2-1 Simple water resources planning flow diagram.

The matrix introduced here identifies two critical differences in underlying modelling assessments that separate existing WRMP practice from the prospective approaches discussed in academia: they measure system performance differently and they treat hydrological uncertainty differently. Measures of system performance can be divided into supply-based metrics, such as yield, and metrics that integrate supply and demand, such as the reliability, resilience and vulnerability metrics proposed by Hashimoto (1982a). Treatments of hydrological uncertainty can be divided into those based on historical recorded data, which in this framework includes perturbed realisations of the recorded data, and those based on synthetic data—i.e., data derived from a stochastic generator. Figure 2-2 depicts a matrix of four distinct water resources modelling approaches defined by these dimensions.

This thesis will argue that there are fundamental characteristics attributable to these dimensions that preclude the need for detailed discussion on the decision-making elements that are particular to a given planning approach. So, for example, the use of a performance measure that integrates supply and demand in the modelling assessment is something that deserves more critical attention than whether the approach recommends a course of action using 'least cost', 'Pareto optimal', 'scenario discovery' or any other decision criterion or visualisation technique. Thus, the thesis is concerned with the fundamental building blocks from which different planning frameworks are fashioned and seeks to define the specific challenges for planning using modelling approaches in these categories. It focuses on the case for shifting the regulated Water Resources Management Planning process away from methods based on "traditional yield analysis" toward methods that employ synthetic flows in stochastic modelling assessments (i.e., "stochastic yield analysis" and "risk analysis").

Hydrological data simulated

Measure of system performance		Historical recorded flows	Synthetic flow replicates
	Supply-based	***TRADITIONAL YIELD ANALYSIS***	***STOCHASTIC YIELD ANALYSIS***
	Integrated	***HISTORICAL BEHAVIOURAL TESTING***	***RISK ANALYSIS***

Figure 2-2 Matrix for characterising water resources modelling approaches.

2.2.2 Characteristics of four distinct water resources modelling approaches

A “supply-based” measure of system performance is taken to be one that derives from a modelling approach that determines the maximum rate of supply that can be sustained by a water resources system subject to a set of assumed conditions and constraints. These assumptions include specific hydrological conditions (input time series), asset capabilities (e.g., storage capacity, maximum pumping rates, maximum flow through pipelines and treatment works, etc.) and failure criteria—usually a maximum permissible storage drawdown, a desired level of service based on water restriction frequencies, and minimum environmental flows and compensatory release requirements where relevant. The resulting metric of supply is known internationally as “system yield” and today in England and Wales as “Deployable Output” (DO).

“Traditional yield analysis”—upper-left quadrant in Figure 2-2—is thereby defined as an assessment of system yield using recorded historical (naturalised) hydrological conditions. There are countless approaches to computing system yield using historical time series, but the subsequent investment planning procedure invariably involves the use of a “supply demand balance”, because the modelling assessment is executed independent of any assumed quantity of measured or forecasted demand. The design paradigm is typically one of capacity expansion: define the least-cost combination of interventions that overcome any supply-demand deficit within a specified planning horizon. A general technical overview of least-cost capacity expansion planning is given in Loucks *et al.* (2005), whist an up-to-date summary of the specific principles followed by water companies in England and Wales, known broadly as the "economics of balancing supply and demand," can be found in Environment Agency (2012) and in more detail in UK Water Industry Research (2012b).

System yield can also be computed using synthetic hydrological data—an approach termed "stochastic yield analysis" in the matrix (upper right quadrant, Figure 2-2). Using this approach, the analyst might generate a large sample of replicate sequences, compute the yield for each and then design the system so that a given percentile of the yield distribution (e.g., 95th percentile) exceeds the demand. Importantly, a stochastic yield analysis can support the traditional forms of decision making and investment planning—namely the least-cost capacity expansion approach described above. To illustrate, Southern Water recently submitted a draft WRMP that makes a case for capacity expansion, and associated infrastructural investment, on the basis of a supply metric derived using a synthetic drought (Southern Water, 2013).

Supply-based performance metrics contrast with metrics that integrate supply and demand. Here the aim of the modelling assessment is not to define the supply that can be sustained by a system, but to observe and quantify the behaviour of the system under particular climate and demand conditions. Performance is normally indicated using thresholds defined either by water storage levels or simulated supply shortfalls. The majority of prospective planning methods that apply these metrics also use synthetic flows, and therefore lie in the lower-right quadrant of Figure 2-2, termed "risk analysis." Synthetic flow replicates tend to be favoured when computing integrated performance measures because the typical aim of such studies is to estimate probabilities for particular consequences—in other words, to determine metrics of risk. This may include likelihoods for specific storage trigger crossings or supply shortfalls of a given magnitude. Since these events tend to be extreme and rare by definition, the brevity of available historical records (typically 50-100 years) prevents reasonable likelihood estimates using these inflow data. Often the historical record is unlikely to contain the severity of event that is of interest. Most systems in England and Wales, for instance, will be designed to easily cope with the worst drought on record and so an estimate of the probability of catastrophic failure would be impossible using a simulation of that record (the simulation would return a probability of zero). Moreover, synthetic flow sequences improve prospects for capturing performance uncertainty, because the stochastic model can be re-sampled to produce alternative plausible realisations (replicates) that differ in structure to the historical sequence.

2.2.3 Research questions

The following specific research questions derive from the definitions set out in Figure 2-2. For concision, the term "stochastic modelling assessments" is used here, and throughout the thesis, to denote the “stochastic yield analysis” and “risk analysis” modelling approaches introduced above.

RQ1. What practical methods could planners deploy to begin exploring the use of stochastic modelling assessments for analysing water resources system performance?

RQ2. What insights might planners gain from running stochastic modelling assessments, and how do “stochastic yield analysis” and “risk analysis” compare in this regard?

RQ3. What practical, conceptual and institutional challenges might water companies face when attempting to plan using stochastic modelling assessments?

2.3 Methodology

The research pursues a qualitative approach based primarily on modelling trials. Whilst each modelling trial is quantitative in nature—comprising mathematical models and statistical analysis—the broader ambition is to identify and characterise the insights made available by, and the challenges associated with, the approaches employed. This is achieved qualitatively, relying on interpretation and practitioner discussion. Each modelling trial can be thought of as a case study in which the subject (or case) is a particular planning framework and the phenomenon under study is the use of the stochastic modelling assessments for exposing vulnerabilities and risks in such a way that might help the planner decide how to invest. Table 2.1 analogises the approach to a more archetypal form of case research in which the phenomenon under study is organisational behaviour and the subjects for case analysis are organisations.

Case studies provide a means to study specific phenomenon in a real-life context and are appropriate for exploratory research and theory building (Eisenhart, 1989; Voss *et al.*, 2002). The purpose is to “generalize to theoretical propositions” by converging findings from different sources (Yin, 1994). So just as the management researcher in the analogy (Table 2.1) need not visit the offices of every firm listed on the London Stock Exchange to build a theory on innovation in large firms, this thesis need not execute every planning framework that employs stochastic modelling assessments to explore the practical and conceptual challenges associated with the use of these approaches for solving real-world planning problems.

The thesis presents four modelling trials sequentially. The approach was iterative, such that the findings from one study shaped the ambitions of the next (as described by Eisenhart, 1989). Each study used a different water resources system, different performance metrics and a different approach to generating synthetic flow data (summarised in Table 2.2). This diversity was both a consequence of the specific challenges that arose for each modelling trial and a deliberate strategy to broaden the range of the research. All water resources systems used were conjunctive-use (i.e., comprising different resource types—surface water, groundwater, desalinated seawater) but dominated by surface water sources. In all cases, the studies were executed using models and software used routinely by water company planners and analysts. Most of the necessary data for executing these studies were obtained from water companies, namely United Utilities and Melbourne Water. These data included water resources system models incorporating asset specifications, abstraction licence and compensation flow requirements and system operating control rules. The named companies also provided catchment inflow records, rainfall records and water demand data (described in detail in each case study chapter). All other data and software used in the studies—such as climate change projections—are available publically. All analyses were conducted using standard desktop computing facilities.

The modelling trial diversity summarised in Table 2.2 prevents fair cross-comparison of methods in terms of the specific modelling outputs, such as the quantified measures of risk. However, since this thesis focuses on more general aspects, such as the form of the outputs, necessary assumptions, data requirements, computational requirements, etc., the variation should not negatively impact the robustness of the overall findings. The individual papers—particularly in Chapters 3 and 5—also report extensively on the generality of the conclusions to water resources planning problems.

Table 2.1 Modelling trials as case studies—analogy with archetypal case research

	Archetypal case research	**This thesis**
Research Question	How do large organisations innovate?	As above – RQ2 and RQ3
Phenomenon under study	Innovation in firms	The use of stochastic modelling assessments in water resources planning frameworks
Case subjects	Large organisations (e.g., Apple, IBM, General Electric)	Planning frameworks that employ stochastic modelling assessments (e.g., “Risk-based planning,” “robust decision making”...)
Entering the field	Approach organisations for access to staff and company documents	Execute the planning framework on real-world system using existing models and data
Data collection	Interview R&D managers, obtain company documents, review archival records, etc.	Record difficulties; note precarious or questionable assumptions required to execute the analysis; discuss process and results with planners.

Table 2.2 Summary of four modelling trials

Thesis chapter	**3**	**4**	**5**	**6**
Planning framework	Risk-based planning	Robust Decision Making	Decision Scaling	Capacity expansion
Water resources system	Ennerdale Water	West Cumbria WRZ	Melbourne Bulk Supply System	Integrated WRZ
Relevant water company	United Utilities	United Utilities	Melbourne Water	United Utilities
System type	Single raised lake with ground water support	Raised natural lakes with ground water support	Impounding reservoirs with desal. support	Integrated multi-basin conjunctive-use
Effective storage	4 GL	7 GL	1800 GL	480 GL
Population served	~50,000	~150,000	~4 million	~7 million
Spatial extent	Single catchment system	Small: Three linked sub-zones	Large: multiple catchments	Very large: multiple basins
Features of note	Short critical period (~90 days); environmentally sensitive area (SAC, SSSI, Habitats Directive).	Weakly interlinked system; highly impacted by environmental constraints on abstraction and releases.	Large over-year capacity (ten years' storage); support from inter-basin transfer and desalination; suffered recent mega-drought.	Substantial source diversity; complex operating rules; strongly impacted by environment constraints (Lake District).
Modelling platform	Aquator	Aquator	eWater Source	Aquator
Code for batch processing	VBA	VBA	R-script	VBA
Synthetic flow data source	UKCP09 Weather Generator	Future Flows Climate ensemble	Multi-site autoregressive flow generator	Multi-site autoregressive rainfall generator
Hydrological model	Basic transform function	AWBM model	N/A	AWBM model
Synthetic replicate sequences used	100 × 100 years (single site)	11 × 150 years (9 sites)	1000 × 100 years (11 sites)	50 × 100 years (23 sites)
Temporal resolution	Daily	Daily	Monthly	Daily

The final fieldwork activity, reported in Chapter 7, comprised in-depth, structured interviews, which were designed to complement the technical analyses of the thesis. These were used to broaden the understanding of practitioner perspectives on the use of stochastic modelling beyond the views of sponsoring companies. The aim was to capture viewpoints on the strengths and weaknesses of existing planning practice and the case for change toward an approach informed by stochastic modelling assessments. The study reports practitioners' perceptions of business risks associated with change to inform discussion on the types of institutional challenge that may not have been easily identified through the modelling trials.

Figure 2-3 describes how the research questions are addressed through the modelling trials and interview approach. Each modelling trial was presented to water resources planners in the organisations concerned. Subsequent discussions informed conclusions on strengths and weaknesses of different planning approaches reported throughout the thesis. The planning frameworks were selected based on contemporary literature, the interests of the project sponsors and the lessons from the prior case studies where applicable. The study in Chapter 6 differs from the others, as it makes no attempt to present an emerging planning framework. Instead, this particular study uses "risk analysis" to address a problem of consistency within the existing planning framework. Nonetheless, useful conclusions can still be drawn from the study because it employs a stochastic modelling approach on a very large and complex system.

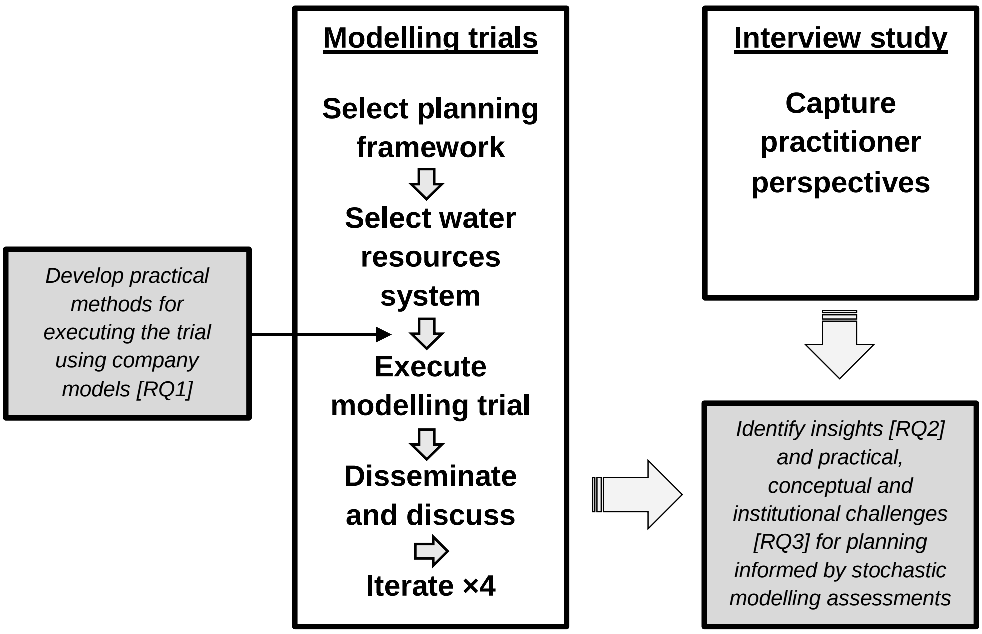
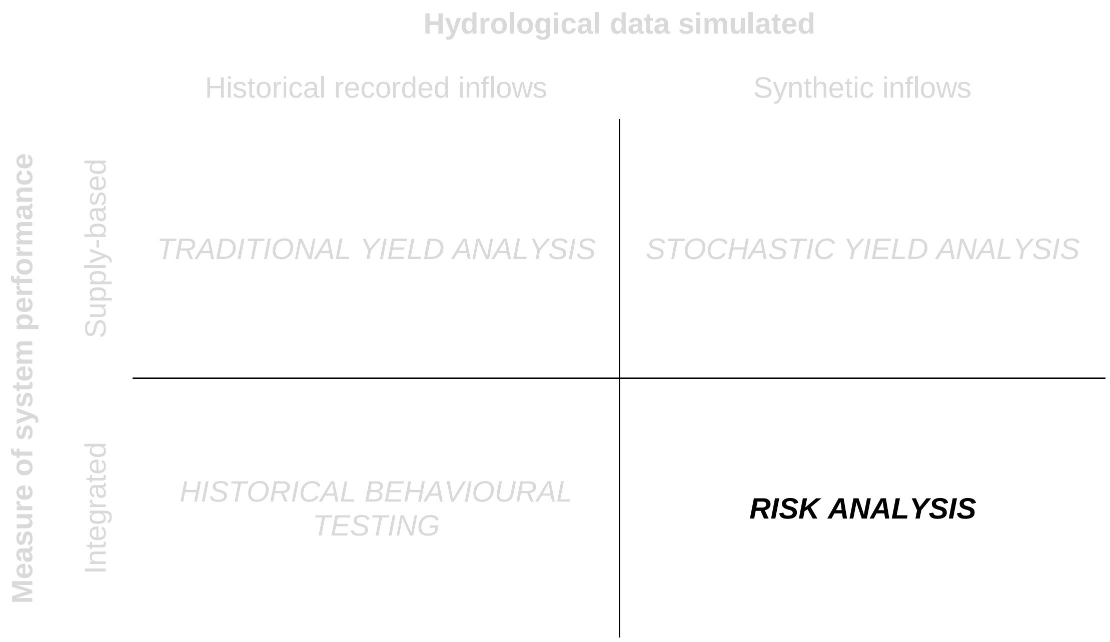


Figure 2-3 Flow chart describing the research framework and the where the research questions are addressed.

CHAPTER 3 RISK-BASED PLANNING

Article title	*Risk-based water resources planning in England and Wales: challenges in execution and implementation*
Co-authors	Richard Blackwell (United Utilities PLC), Mark Smith (United Utilities PLC), Paul Jeffrey (Cranfield University).
Co-author contributions	Raw data provision, style guidance and corrections, general discussion, provision of variants from UKCP09 Latin Hypercube Sampling study.
Publication status	Published in *Urban Water Journal*.
Reasoning for case and site selection	The "risk-based" approach had been advocated in a then recent publication (Hall *et al.*, 2012a) and discussed in an UKWIR study (UK Water Industry Research, 2012a). Ennerdale Water was an appropriately simple system to begin investigating stochastic modelling using Aquator.



3.1 Introduction

Planning the efficient and sustainable use of water resources involves both great skill and rigorous analysis. In order to design reliable, cost-effective supply systems, planners must attempt to understand how the natural availability of water, and society's demand for it, will change in future. The potential for change in drought frequency and severity requires particular attention. Yet many of the mechanisms that influence hydro-climatic extremes are poorly understood (Koutsoyiannis *et al.*, 2009) and so robust design must somehow accommodate uncertainty.

Hydrologists and planners have traditionally dealt with uncertainty through precautionary measures, such as reserve storage in reservoirs (Salas, 2013). Stakhiv (2011) suggests that these design principles have served society well, but that new approaches are needed for dealing with the uncertainty associated with human-induced climate change. Importantly, Global Climate Models (GCMs) have failed to adequately simulate observed precipitation patterns under historic emissions (Kundzewicz *et al.*, 2008; Blöschl and Montanari, 2010) and therefore offer limited assistance to water managers that seek to predict future drought characteristics. Countering the view that human-induced climate change has fundamentally shifted the planning problem, a number of recent contributions have emphasized the Hurst-Kolmogorov phenomenon to draw attention to the naturally complex and unpredictable nature of hydro-meteorological systems (Koutsoyiannis *et al.*, 2007; Koutsoyiannis, 2011; Lins and Cohn, 2011). Nonetheless, the climate change adaptation agenda has focused new attention on uncertainty and presented the water resources community with a range of probabilistic climate information (predominantly from GCMs), fostering new dialogue around the tenancy of deterministic metrics in contemporary water resources planning (Milly *et al.*, 2008; Salas *et al.*, 2013). In England and Wales, Hall *et al.* (2012a) and Hall and Borgomeo (2013) have brought this debate sharply into focus through their strong advocacy for a risk-based planning framework informed by a probabilistic approach to water resources system assessment.

Tasked with looking to the future of water resources planning in England and Wales, a recent UK Water Industry Research study found that 'numerous scientific papers illustrate components of such a [risk-based] framework although none address the practical challenges that would be faced by the water companies' (UK Water Industry Research, 2012a). Previous contributions have developed the application of probabilistic climate projections to natural stream flows (e.g., New *et al.*, 2007; Manning *et al.*, 2009) and water resources system models (e.g., Lopez *et al.*, 2009). UK Water Industry Research (2012a) now note a need to advance practical methodologies for: characterising supply and demand uncertainties; linking the probabilistic outputs of (e.g.) the UK Climate Impacts Programme to water resource system models; estimating probabilities of supply-demand deficits; and optimising water resources system design by weighing investment costs against benefits in terms of reduced risk. These intentions reflect a sector-wide ambition to develop risk-based planning methods for dealing with climate change and hydrological uncertainty (Kundzewicz *et al.*, 2008; Brown and Baroang, 2011; Salas *et al.*, 2013).

In this paper we describe the execution of a risk-based analysis on a simple, small-scale case study, which we term “the Ennerdale Problem.” We compare outcomes against those that would be generated using a more conventional analysis, providing a first indication of the implications of using a risk-based framework in England and Wales. We aim to demonstrate how to practically use—and benefit from—the risk-based methodology, and identify challenges for implementing this type of approach from the viewpoint of a private regulated water service provider.

3.2 Risk and uncertainty in water resources planning

The risk-based methodology—as we describe it—attempts to remedy two fundamental weaknesses of contemporary water resources planning practice. The first relates to a reliance on limited historical hydrological records as the sole model input to simulate future droughts; the other is the necessarily simplistic assessment and design metrics that dominate as a result of having to rely on historical records. The risk-based approach relies on stochastic models to extend available climate information into a large number of synthetic replicate sequences. The enlarged sample size enables assessment and decision-making based on metrics of probability and consequence.

Stochastic hydrology first emerged in the 1960’s as a major output of the Harvard Water Program. Maass and Hufscmidt (1959) noted that ‘the Harvard Program proceeds on the assumption that reliance on stream gauging records alone is a weak link in the present methodology’ (in Reuss, 2003). The Harvard Water Program investigated the use of simulations on (then) high speed computers and produced quantitative models of stationary stochastic hydrology that would support more thorough assessments of ‘reliability’—the probability of system failure over a given time period (Note that according to Fiering (1997), stochastic hydrology was developed as a means to identify weak points in large interconnected reservoir systems, rather than to conduct risk-based system assessments). These advances spawned new ideas around water resource system assessment; a widely cited contribution by Hashimoto *et al.* (1982a) described a method based on reliability (as described above) and two additional system performance metrics (resilience and vulnerability) that would seemingly provide a thorough description of risk and a sensible grounding for rational management decisions in operating and planning. Whether any of this research seriously influenced water resources planning practice is debatable. Looking back on the progression of his early path-making developments at the Harvard Water Program, Myron Fiering lamented that ‘[stochastic hydrology] really has downed no enemy planes, nor has it sunk any enemy subs. What is worst, it has only rarely been fired in anger; it has engaged in lots of target practice, but rarely in combat' (Fiering, 1997). In England and Wales, the recent progression from simulating historical flows to simulating perturbed versions of those flows hardly represents a shift toward a stochastic exploration of water resources system performance. More broadly, the standard capabilities of well-established commercial and academic water resources modelling software packages (e.g., WEAP, MIKE BASIN, HEC-ResSim, RIBASIM) would suggest that non-probabilistic scenario simulation remains the dominant method for exploring future system performance. In agreement with this view, Brown and Baroang (2011) noted that 'research has largely proceeded from a scenario-based rather than a risk-based framework of analysis.'

Given this lack of previous uptake, why might risk-based methodologies now displace the existing methods, or at least occupy a more prominent role in practice? De Neufville (2004) offered two possible reasons. First, the increase in analytic speed and memory capacity of desktop computers has empowered the mass of skilled practitioners to run thousands of system simulations with ever decreasing time and cost (e.g., Asefa *et al.*, 2014). Second, methodological advances over the last two decades may allow analysts to use the probabilistic outputs of risk-based analyses to greater effect. 'Real options analysis' and 'robust design' are identified as revolutionary developments for 'drawing meaning out of these [probabilistic] calculations so we can use the results in productive ways.'

Both factors noted above seem to bring risk-based methodologies within reach, but neither could be considered a driver for change in practice. Rather, the increasing prominence of uncertainty, brought about by the recognition of uncertain climate change impacts on extreme weather events, has driven recent developments in stochastic hydrology and water resources management. Two decades of climate modelling research has culminated in probabilistic climate projections and, in the UK, user-friendly tools for generating synthetic hydrological time series that represent a range of climate model projections (UKCP09 weather generator; 'Future Flows Climate' – Prudhomme *et al.*, 2012). In this study, we execute a risk-based analytical method that employs these data, and expose the practical problems that would arise if such an approach were industry-standard.

3.3 Case problem

3.3.1 Background—the Ennerdale problem

The West Cumbria Resource Zone (Figure 3-1) is managed by United Utilities PLC (the incumbent water company) and supplies water to approximately 150,000 people in northwest England (annual average demand ~ 49 ML/d). Ennerdale Water and Crummock Water provide the majority of water storage in the zone and supply the towns of Whitehaven and Workington respectively. These are artificially raised natural lakes, which are located in relatively wet catchments (annual precipitation ~ 1800 mm) characterised by steep rocky terrain and associated rapid rainfall-runoff response. Other water sources in the zone include a combination of small reservoirs, boreholes and stream abstractions. Connectivity within the zone is relatively weak: the network can transfer about a quarter of the Ennerdale-fed demand from Crummock Water (and vice versa) and is set up to do so only during drought or emergency situations.

The viability of Ennerdale Water as a supply source has come under increasing scrutiny in line with legislated measures to protect and enhance the local aquatic environment. Ennerdale itself is a designated Site of Special Scientific Interest (SSSI), and the River Ehen—for which Ennerdale is the natural source—is both an SSSI and a designated Special Area of Conservation (SAC) under the Habitats Directive (EC, 1992). These designations have both extended and strengthened the environmental thresholds and consents at Ennerdale.

In the previous water resources planning cycle (2005-2010) the environmental regulator mandated an approximate 50 % increase in the flows apportioned from Ennerdale to the River Ehen. The aim was to protect freshwater pearl mussels—an endangered and protected species. The water company responded by planning a new groundwater scheme to offset any impact on security of water supply for its customers. More recently, however, the environmental regulator indicated that further increases in compensation flow would be needed to conserve river ecology in the long term. These changes have threatened the fine balance between meeting customer supply whilst providing adequate compensation water to the River Ehen and avoiding environmentally unsafe lake drawdown. But weighing hard against any compulsion to abandon Ennerdale is the financial cost of mobilising alternative water resources in the region (estimated £150-500 million depending on solution).

3.3.2 Framing the Ennerdale problem

We wish to emphasise that various rapidly changing complexities obscure the problems at Ennerdale beyond the simplistic understanding presented in this paper. We have deliberately excluded these from our analysis in order to clearly communicate both the problem—which now becomes hypothetical—and the implications of our results. The detail should not be compared to United Utilities' draft Water Resources Management Plan (United Utilities PLC, 2013), which makes additional allowances for various aspects excluded from our analysis. The Ennerdale problem serves as a vehicle with which to: demonstrate a practical method of executing the risk-based analysis; understand the benefits of this approach compared with a conventional analysis; and identify challenges for industry-wide implementation. Thus we are concerned solely with the approach and principles of deriving and articulating useful risk-based metrics of system performance, rather than commenting on specific details of this case.

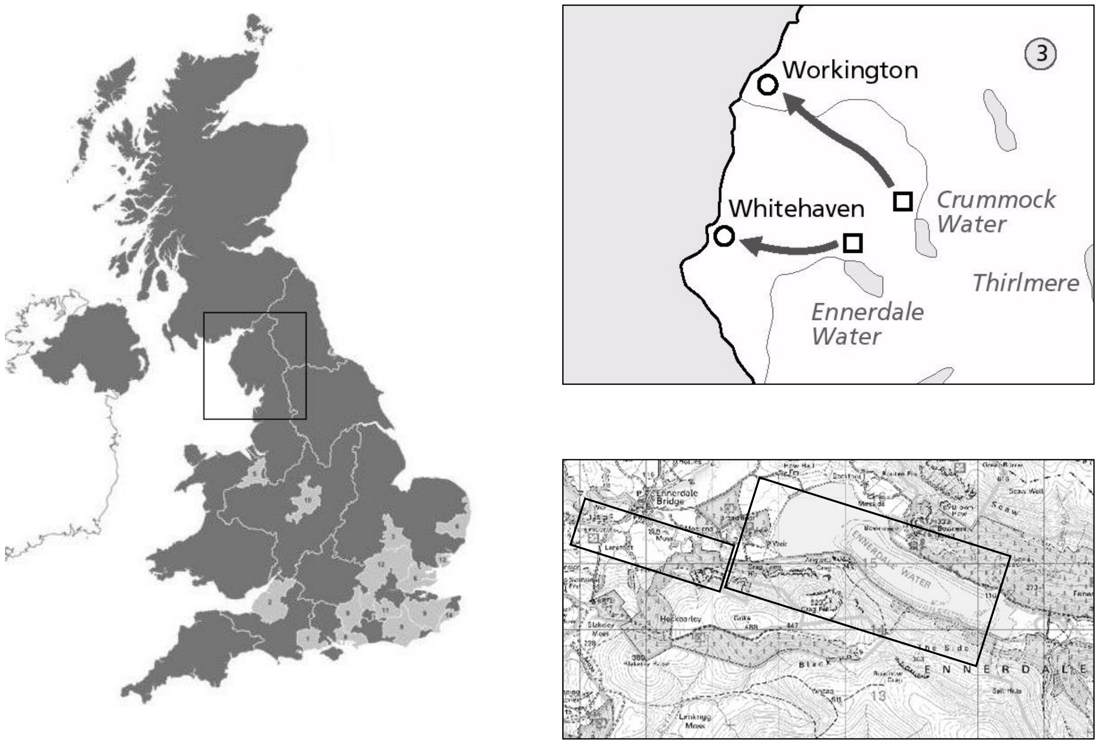


Figure 3-1 West Cumbria resource zone. Left: location in UK water company regional map. Top-right: West Cumbria water resource zone predominant transfers (black outlined square represents major water treatment works, solid line represents gravity supply pipe). Bottom right: OS map showing Ennerdale Water and River Ehen.

Our case problem is conceptualised as a simple, single-reservoir system (Figure 3-2). Ennerdale Water supplies water directly to the River Ehen through mandatory level-dependent controlled compensation releases, or as spillage when the lake rises above weir crest. The water company is permitted to use approximately 5000 ML storage (the top 1.7m of a 40m deep lake—although drought powers are required to legally take water beyond 1.35m depth) to meet both the compensation flow requirements and an average customer demand of 28.1 ML/d. We assume that the South Egremont borehole (annual licence volume = 2000ML) will supply 6.4ML/d to Ennerdale Water Treatment Works whenever the lake level drops below weir crest, thereby offsetting the same demand from Ennerdale Water. Operational triggers (defined in Figure 3-3) indicate when drought management measures would be initiated in line with lowering lake levels. We assume that the system is a discrete, self-contained water resources system (in reality the demand served by Ennerdale can be augmented by supply from adjacent reservoirs). The case problem exhibits low-probability, high-consequence risks (e.g., running out of water) and the specific geographical/socio-political situation severely limits the infrastructural and operational intervention options.

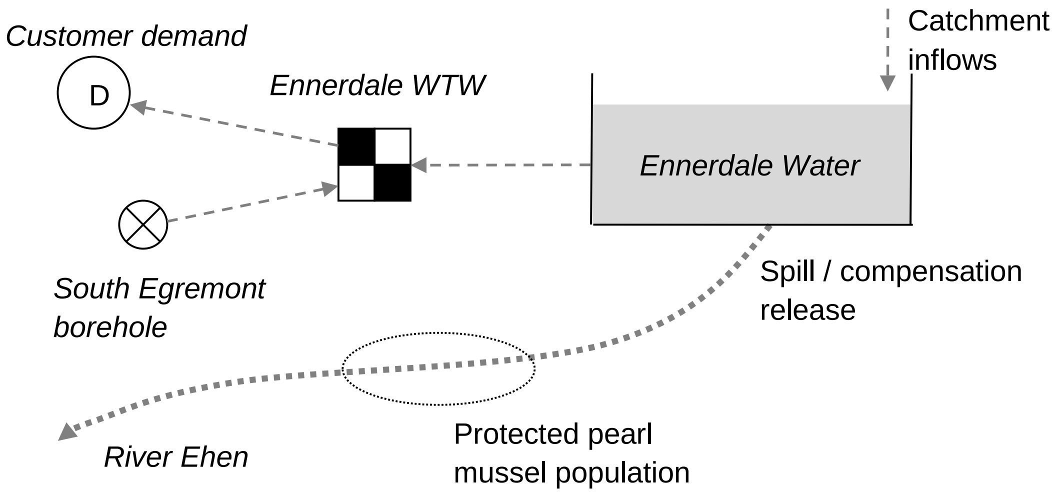


Figure 3-2 Schematic of the Ennerdale supply system

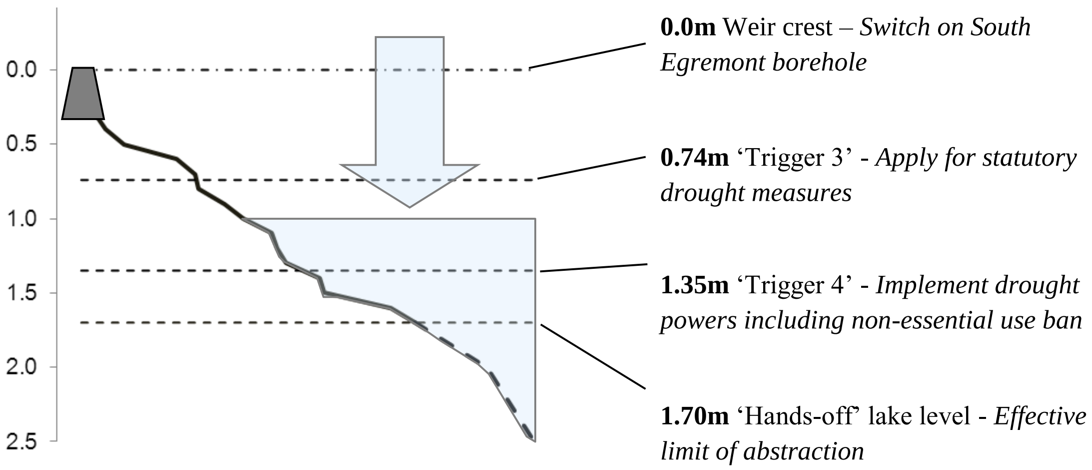


Figure 3-3 Reservoir profile diagram (features not to scale) highlighting key thresholds.

3.4 Conventional methods for dealing with uncertainty

3.4.1 Deployable Output and the Ennerdale problem

Figure 3-4 shows the output of a conventional water resource system assessment (Environment Agency, 2012) applied to the Ennerdale problem. The supply metric—"Deployable Output"—is defined as the greatest customer demand the system can supply without incurring failure during full simulation of a 50-year historical reservoir inflow series (1961-2010). Failure occurs at a specified lake level defined using a reserve volume of 30 days' demand lying above the hands-off lake level. This is essentially an elaborated version of the Rippl (1883) mass curve reservoir sizing method with the addition of a reserve volume (termed "Emergency Provision" in England and Wales), which is common practice across the water resources planning profession in England and Wales (Ratnayaka *et al*., 2009), the United States (Water Research Foundation, 2013), Australia (Erlanger and Neal, 2005) and elsewhere.

Two elements of the output in Figure 3-4 are worthy of note: an almost 50 % drop in Deployable Output (the supply metric) occurring five years into the planning period, which corresponds with the indicative future compensation flow increase described above, and the apparent supply-demand deficit opened up by these changes. The planning process is designed such that a deficit exacts a response by way of infrastructural investment to ensure the metric of supply meets that of demand plus an additional allowance for uncertainty, known as "Headroom".

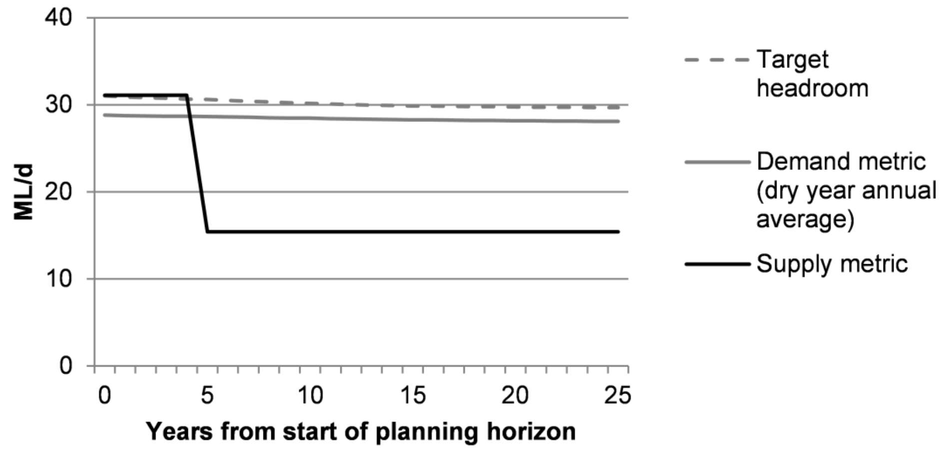


Figure 3-4 Output from a conventional analysis for the Ennerdale problem.

3.4.2 Weaknesses in the conventional approach

This conventional method of analysis is fundamentally weak; it lacks transparency and, more worryingly, it (unintentionally) misleads. Specifically, it perverts our indeterminate understanding of future conditions by fixing uncertainty into precise projections and concrete language (i.e., 'deficit'). We can bring these weaknesses to light by considering how the method deals with uncertainty.

The method embodies two distinct allowances for uncertainty. Firstly, Headroom combines, in a probability distribution around the demand, the uncertainties arising from demand forecasting error, climate change impacts on Deployable Output (assessed through scenario testing), capabilities of new assets and data error. Hall *et al.* (2012a) provide a thorough critique of this approach, noting that the percentile of headroom used to define a tolerable level of risk ("Target Headroom") merely offers up the possibility of a pseudo risk assessment because it matches probability to the abstract metric of Deployable Output rather than tangible consequences (e.g., impacts of drought measures, supply failure). We agree with this critique, and offer an additional component to the debate by considering a second, often overlooked, allowance for uncertainty: that of 'Emergency Provision'.

Emergency Provision (more informally known as Emergency Storage) deals with the uncertainty of future flow variability. Conventional analysis justly recognises that a historical flow sequence fails to fully represent the range of possible droughts that might occur in future. The Emergency Provision protects against the risk of those events' occurrence and thus forms a mandatory design component for water resource systems in England and Wales. It is represented as a volume of reservoir water effectively removed from the modelled storage to induce earlier failure in simulated historical droughts. This type of measure is unavoidable in an analysis that relies on a single simulation of the historical inflow record. We can clearly see why it was introduced in England and Wales on the back of the 1995/96 droughts that posed serious questions of water resource system reliability across northern England (Department of Environment, 1996; Smithers and Walker, 1997; Walker and Smithers, 1998).

Yet if we are to understand water resources system risk then we must acknowledge the existence of Emergency Provision and its influence on the problem at hand. By sizing Emergency Provision subjectively, water planners in England and Wales (perhaps unwittingly) impose a largely undefined assumption regarding the level of risk against which their systems are designed. They can choose to apply an Emergency Provision of between 15 and 45 days' demand volume (UK Water Industry Research, 2012b), but planning guidelines fall short of mandating a demonstrated understanding of the drought probabilities these volumes protect against (1 in 100, 1 in 1000, 1 in 10,000 year...?) This means that a Deployable Output that meets target headroom—say, 95% of total headroom uncertainty—seems robust to deal with 95% of plausible futures without "failure". But because that "failure" relates to a drought of undefined probability of occurrence, the whole analysis becomes ambiguous and potentially misleading.

Using the Ennerdale problem to demonstrate our point, had we decided to subjectively choose a 15-day emergency provision (as opposed to the 30-day industry standard) our future system would yield 27.5 ML/d during the worst historical drought (as opposed to 15.4 ML/d, projected in Figure 3-4). In other words, this subjective assumption can impose on our results the difference between a system on a knife edge, necessitating abandonment and associated investment in new supply options, and one exhibiting a minor shortfall that could, according to the analysis, be easily tempered with relatively small infrastructural or operational changes.

Adding to this critique, we would note that this Emergency Provision is based on assumed demands (UK Water Industry Research, 2012b) and thus overlooks the possibility of variability across different reservoirs with regards to inflows from catchments and network connections during drought. So two separate resource zones with the same emergency provision (in days' demand) might provide buffers against drought events with radically different probabilities of occurrence. The Emergency Provision produces an analysis that is both opaque in its definition of risk and biased as a method for comparing risks across different resource zones.

3.5 Execution of a risk-based analysis

To contrast the conventional analysis reported above with a risk-based planning framework, we describe the execution of an analysis that draws on the ideas of Hall *et al.* (201a2) and Brown and Baroang (2011). Interested readers should refer to these articles for a full description of the principles of a risk-based water resources planning methodology. We translated these principles into four broad objectives: (i) to integrate supply and demand uncertainties into the water resources modelling procedure; (ii) to predict the probabilities of meeting levels of service related to tangible consequences for customers (supply restriction), companies (e.g. operational, legal and reputational costs of supply disruption) and the environment (e.g. failure to meet compensation flow requirements); (iii) to consider impacts arising from droughts of different severity and duration; and (iv) to understand risk reduction as a goal to be weighed against the costs of system improvement.

Our methodology comprises a number of formal steps, detailed in the following subsections. We began by defining the uncertainties to be integrated into the modelling analysis. Supply uncertainty was modelled using 100 synthetic reservoir inflow series derived from the UKCP09 Weather Generator (a tool that produces daily stochastic precipitation and evaporation time series under different climate model realisations). This number of series was judged adequate to capture the uncertainty without being too large to impose impractical run times. Demand uncertainty was sampled using Monte-Carlo simulation from a probability distribution function defined using scenario analysis. The water resources system was modelled in Aquator (Oxford Scientific Software, 2008) and controlled externally using VBA code to automate multiple runs, each time injecting sampled model parameters (i.e., randomly sampled demand) and extracting output. We captured extended time series data on lake storage levels and analysed for threshold-crossings corresponding to operational responses that imply tangible consequences. Finally, we used these results to inform a more nuanced understanding of risk.

3.5.1 Supply uncertainty—synthetic inflow series

The UKCP09 Weather Generator produces—for any of 10,000 climate future variants across three emission scenarios—plausible daily weather series using a stochastic model (Kilsby *et al.*, 2007). Long stationary time series are available for 5km grid squares, spatially correspondent to UK Met Office historical gridded weather data (and calibrated against this data). We were able to corroborate this gridded output as meteorologically coherent with the Ennerdale catchment by comparing historical monthly gridded rainfall totals with monthly rainfall totals derived from the arithmetic mean of rain gauge data from within the catchment.

A representative spread of 100 from 10,000 UKCP09 model variants was identified using Latin Hypercube Sampling of the standard UKCP09 output. The Weather Generator was then used to produce 100-year stationary weather series for each of these 100 variants at the medium (SRES A1B) emissions scenario for the 2030s time slice (i.e., projected climate conditions for the time period 2020-2049). Next, the weather patterns were routed through a simple rainfall-runoff transform function. The function was calibrated and validated using a 45-year historic daily precipitation and associated modelled potential evapotranspiration (PET) series together with corresponding historic daily reservoir inflows. Further validation was achieved by comparing reservoir yield across different drought years (measured versus modelled). The most severe simulated droughts yielded within 2 – 5 % of the reservoir yields for the same droughts in the measured series. This evidence, in combination with a Nash-Sutcliffe efficiency score of 0.695 for the full record, was deemed adequate to accept the model.

3.5.2 Demand forecast uncertainty

This section describes how demand forecast uncertainty was characterised and implemented in our risk-based analysis. A more detailed account of how to derive a demand forecast uncertainty distribution lies outside the scope of this paper and is covered elsewhere (e.g. Dziegielewski and Baumann, 2011; Environment Agency, 2012). Our primary intention here is to reflect on what the demand forecast represents and describe how the uncertainties around it can form an integral part of a probabilistic water resources modelling procedure.

Demand forecast uncertainty was characterised using a probability density function. Distribution type and parameters were defined using evidence from four alternative scenarios applied to household water demand micro-components (e.g. shower, washing machine, garden sprinkler etc.). Scenario definition and influence over the micro-component parameters (i.e. ownership level; volume of water per use; frequency of use within households) followed the guidance of Westcott (2004) and Memon and Butler (2006) respectively. Base year micro-component data were drawn from recent ownership surveys conducted on behalf of the water company and augmented with prior research into average UK volumes/frequency of water use associated with different components (Chambers *et al.,* 2005). Non-household demands were also forecasted and the influence of leakage and metering were excluded as elements of uncertainty because these aspects could ultimately form potential intervention options for alleviating drought risks.

The uncertainties in this element of the analysis relate to population growth, economic growth, societal and behavioural changes, technological advances and various other largely indeterministic—that is, impossible to accurately forecast—trends. Thus we saw no logic in setting up a dependency between the sampling of this distribution in accordance with the climate future being modelled in any given simulation. However, for any one of these randomly selected indeterministic futures, the volatile inter- and intra-year demand time series would draw strong dependency from weather conditions. For instance, dry hot summers drives up demand as people turn on garden sprinklers, and extremely cold winters correlate with higher demand caused by increased pipe burst frequency. We excluded these dependencies from our analysis for simplicity, but note that this limitation could be avoided by calibrating a model to link inflow series to demand profile based on measured demand and temperature data.

A mean dry-year annual average demand of 28.1ML/d dry-year annual average demand was adopted on the basis of the scenario-based micro-component analysis. This represented the end of the 25-year planning horizon (2039/40) and was complemented with a symmetrical triangular uncertainty distribution with limits of ±9% (based on the micro-component scenario analysis described above). The distribution was coded into a VBA macro for automatic random sampling on the inception of each simulation of the water resources system.

3.5.3 Water resources system simulation and output analysis

The Ennerdale problem and assumed operating conditions (compensatory releases, reservoir spill, borehole triggering etc.) were modelled using Aquator—a commercial mass balance simulator for water resources systems (Oxford Scientific Software, 2008). We designed and implemented a VBA macro procedure to effect external control over the model. The code included the randomised sampler to select demand from our pre-defined uncertainty distribution. It then selected (by order of sequence) one of the 100 alternative flow series and implemented it in the model along with the randomly sampled demand value before initiating a model run for the full 100-year flow series. On completion of the model run, the code extracted the daily calculated lake storage levels for the full run length (100 years) and pasted these data into a spread sheet for post-modelling analysis. The code then repeated the process for each of the remaining synthetic inflow series. The output of the procedure comprised 100 series of daily lake levels, each 100 years in length. Annual occurrence statistics were extracted for three alternative levels corresponding to the thresholds identified in Figure 3-3.

3.5.4 Results

Recall our primary aims for this study: to discuss the practical, beneficial applications of a risk-based methodology and to identify implementation challenges from the viewpoint of a private regulated water service provider. The Ennerdale problem serves as our vehicle to realise and communicate these methods and conclusions. Thus, the following analysis focuses on output type and quality, and deliberately skims over the fine detail of the case study.

Equipped with 100 replicate sequences of daily reservoir storage data, we were able to generate histograms for threshold-crossing drought events of given severity (i.e., specific lake level threshold) and duration (i.e., length of time the reservoir level lies below that threshold). Figure 3-5 displays eight histograms representing the spread of event frequencies across the 100 alternative simulated future climate and demand scenarios. These represent the ‘Trigger 3’ threshold (i.e., apply for statutory drought measures) for event durations that meet or exceed 1, 3, 7, 10, 14, 21, 28 and 35 days. To clarify – these durations represent time periods during which the lake level drops below the threshold of interest. Longer spells of dry weather—which draw down the lake toward the given threshold—would necessarily precede these events in all cases. The histograms in Figure 3-5 display an increase in mean return period with increasing event duration. To illustrate, the 1-day event mean return period = 3 years approx. (30-40 counts in the average 100 year series) and the 35-day event mean return period = 20 years approx. (5 counts).

We also have enough information to attach confidence levels—using the spread of results—for meeting, exceeding or failing to meet given levels of service. Figure 3-6 shows how we were able to extract probabilities of failing to meet different service levels from our histograms, this time using the 1.7m ‘hands-off’ threshold (note that we have converted the x-axis from number of failures to the corresponding level of service—e.g., two trigger crossings in the 100-year sequence signifies a 1 in 50 year level of service for that particular threshold). Drawing these together in a single diagram, we derived a snapshot of system performance that maps probability of failure to meet given levels of service against event durations. Note that the analysis can be read in two ways; a 5% chance of failing to meet a particular level of service corresponds to a 95% chance that that level of service would be met or exceeded. We extend this output in a risk assessment outlined in Figure 3-7 to tie in tangible consequences for each event type.

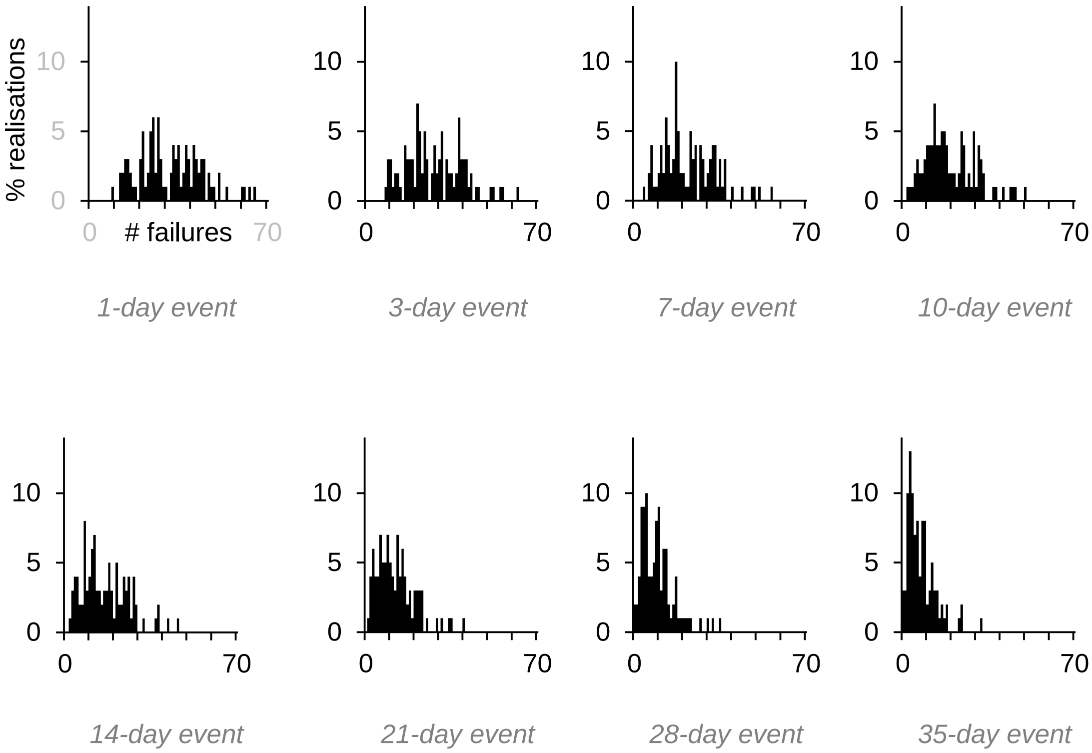


Figure 3-5 Histograms for 0.74m ('trigger 3') threshold for number of counts of event of given minimum duration (e.g. 7-day event) in the 100-year synthetic series. Y-axis can be read as number of series with x counts; or % of series with x counts (since we simulated exactly 100 series).

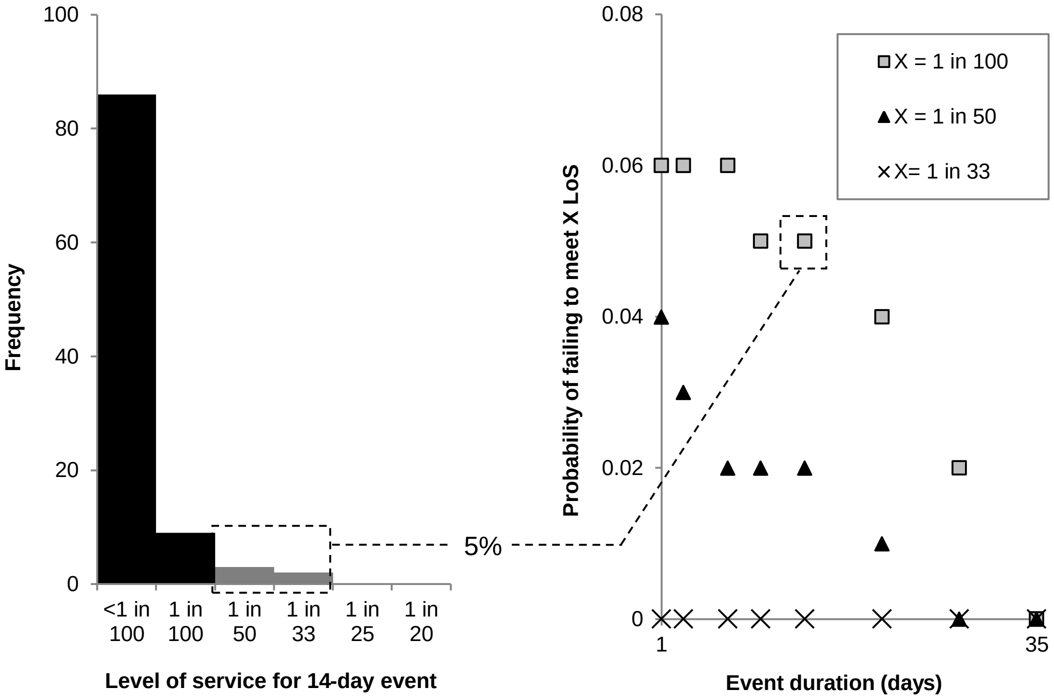
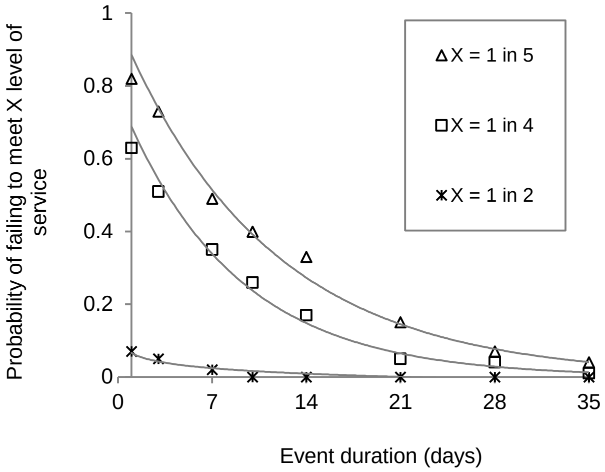


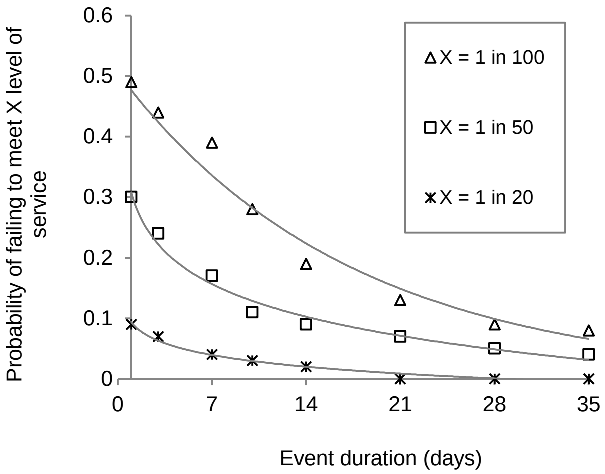
Figure 3-6 Left: histogram for 14-day duration events at the 1.70m 'hands-off' threshold (limit of abstraction). 'Frequency' can be read as number or percentage of realisations representing the given service level. The histogram is separated into dark and light bars to highlight a 5% probability of failing to meet the 1 in 100 level of service for this event. Right: Event-duration probability profiles for three levels of service – highlighting how these profiles are produced from histograms.



a. "Trigger 3"—0.74m drawdown

Event: Apply for drought permit/order and statutory non-essential use ban**.**

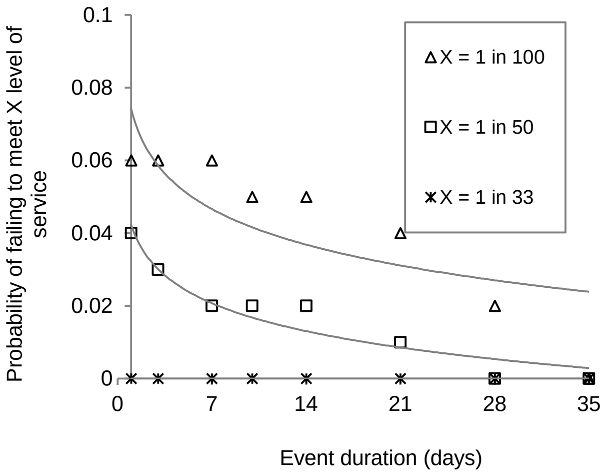
Tangible consequences: Costs of applying for drought planning order applications; costs of implementing demand management actions – e.g. public water efficiency campaigns; costs of bringing more expensive resources online (applicable at weir crest in this case).



b. "Trigger 4"—1.35m drawdown

Event: Implement statutory measures, including non-essential use ban**.**

Tangible consequences: Customer inconvenience (non-essential use ban), company reputational damage; visual impact of lake drawdown and possible adverse impacts on lake ecology.



c. "Trigger 4"—1.70m drawdown

Event: Limit of abstraction breached**.**

Tangible consequences: Severe service restrictions (e.g., stand pipes and rota cuts); tankering of water (financial cost and road disruption); severe company reputational damage; adverse impact on lake and river ecology.

Figure 3-7 Ennerdale risk assessment for the 2020-2049 (2030s) time slice.

3.6 Discussion

Table 3.1 compares key features of our risk-based approach against the existing assessment methodology used in England and Wales. The output decision metrics exhibit the most important distinction—we term them as simple and complex respectively. Undoubtedly, the conventional methodology arms planners with clear-cut output and constitutes a powerful tool for communicating system risk amongst stakeholders (Figure 3-4 exemplifies this). Yet, as we previously demonstrated, the ambiguity around these simplistic metrics leads to biased and potentially misleading descriptions of system risk. The risk-based analysis attempts to justify its complexity—and the inherent difficulties in producing, dealing with, and communicating the complex outputs—by offering a stronger footing for rational decision making.

Table 3.1 Key features of conventional Deployable Output (DO) methodology and risk-based method as described in this manuscript

Feature	**Conventional DO method**	**Risk-based method**
Hydrological model input	Historical record (typically 50-100 years’)	Ensemble of synthetic flow time series
Climate change uncertainty	Quantified as an impact on DO by perturbing the historical record; added to Headroom	Integrated through multiple simulations of the resource system using stochastically-derived flow data
Flow variability and extreme event uncertainty	Accommodated through a subjectively defined Emergency Provision	Integrated through multiple simulations of the resource system using stochastically-derived flow data.
Other uncertainties (e.g. demand)	Assessed in terms of impact on DO, defined as probability distributions and combined in the ‘headroom’ buffer	Integrated through multiple simulations of the resource system and Monte-Carlo sampling from a pre-derived uncertainty distribution
Output decision metrics	Simple - binary understanding of system performance based on whether DO exceeds demand (+ target headroom)	Complex - various potential impacts defined by probability and consequence

The risk-based method described above offers two advances in the discussion of how to execute and benefit from assessment methods that deploy ensembles of hydrological model input. First, we produced probabilities of tangible consequences (e.g., severe supply restrictions), rather than probabilities of 'deployable output', which are loaded with a hidden, subjective safety margin. In contrast to previous UK-based climate impact uncertainty studies (e.g., Dessai and Hulme, 2007), we thereby provide a clear view of both system performance and the modelling assumptions entailed. Second, we extended the analytical understanding of consequence in risk-based planning methodology by examining probabilities of events of varying duration. This could open up the possibility of valuing those consequences, a practical impossibility when all events (whether days, weeks, months…) are considered in binary terms as either failure-inducing or not.

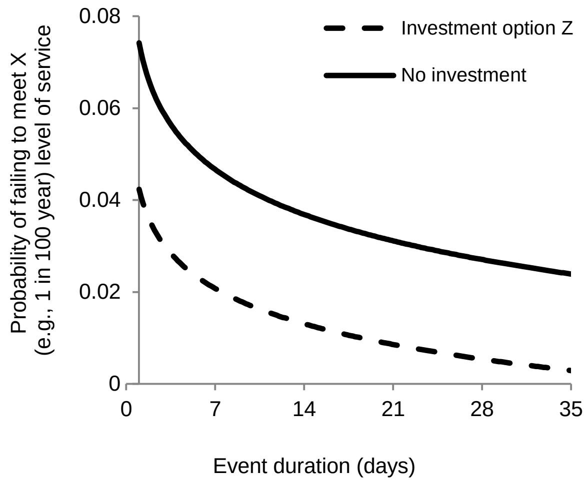
In our simple system, we found the risk-based analysis both practically feasible and beneficial as an advance on the conventional analytical method. Below we support this conclusion by outlining two specific advantages of the approach. As our work falls short of unequivocally demonstrating the pertinence of this approach as the basis for a new water resources planning paradigm, we also identify several nontrivial challenges for scaling it up.

3.6.1 Advantages of a risk-based analysis

Figure 3-8 provides two practical examples of how the risk-based method described above might improve our understanding of water resource system performance. First, the risk-based analysis could inform a more rational and transparent options appraisal. Current economic appraisal methodology defines the least-cost combination of investments to ensure demand is met by deployable output, which, as we have seen, rests on poorly understood safety margins. The probabilistic analysis of tangible consequences offers up a more comprehensive assessment of risk that might allow analysts to attach values to the outcomes of alternative interventions. We wish to avoid overstating this benefit; below we outline how further research and clear guidance will be needed to make best use of probabilistic projections to reach rational investment decisions. Nevertheless, we see potential for a different form of options appraisal using the database of probabilistic output. For instance, multiple time series of reservoir drawdown scenarios could inform a decision analysis based on robustness criteria (described by Lempert *et al.*, 2003) by quantifying the value of 'regret' associated with each strategy compared across different futures.

Second, the approach should enable the analyst to compare the risks attached to different water resource zones. In many cases, conventional analysis fails to compare resource zones fairly because the emergency provision assumption neglects likelihood of reservoir inflows from catchments and interconnections during drought. Applying the risk-based method to two resource zones that exhibit similar risk levels in a

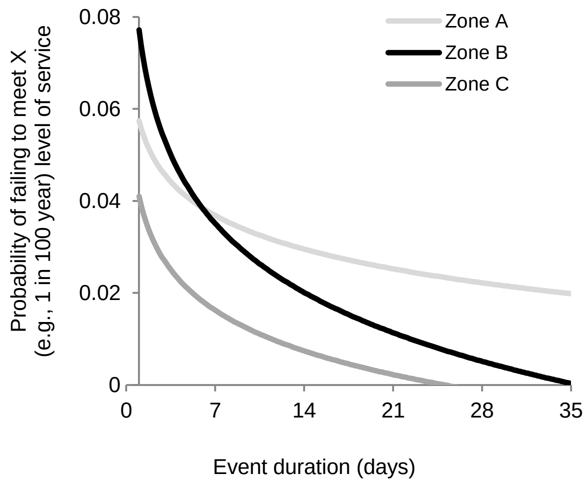
Deployable Output analysis might unveil previously unnoticed discrepancies to drive a more targeted investment strategy. Such a comparison—if applied to relatively small resource zones—may be feasible without further development of the methodology presented in this paper. Of course, the emergency storage assumption does not necessarily preclude fair comparison within the conventional planning framework. We believe analysts could size up the emergency storage in each zone in such a way that corrects for bias. However, a review of the 2013 draft Water Resources Management Plans (WRMPs) issued by water utilities in England and Wales exposes only one attempt to correct for bias. As yet, there appears to be no documented method for dealing with the emergency storage bias within the existing framework.



a. Compare investment options

Analysis description: Measure the benefits of different investment options in terms of reduction in risk.

Benefits to planners: Investment benefits measurable in terms of impact on tangible consequences and their probabilities of occurrence; improved prospects for weighing costs of investment against the benefits in terms of risk reduced.



b. Compare Water Resource Zones

Analysis description: Compare different water resource zones – without emergency provision bias - for risk of failing to meet X level of service.

Benefits to planners: Unbiased comparisons of different zones, leading to more rational and targeted investment strategy; option to tie in other comparative risk measures—e.g. population affected; vulnerability of environment and human populations in zone under study—leading to fuller picture of real risks.

Figure 3-8 Examples of possible beneficial applications of the risk-based methodology

3.6.2 Technical execution practicalities

We anticipate a number of criticisms relating to the technical limitations of the approach described here. For example: the parameters and shape of a future demand uncertainty distribution were not (nor can they be) validated; an intra-annual demand profile (with dependencies mapped to modelled weather conditions) was not included; the climate projections were based solely on a 'medium' emissions scenario; the appropriateness of the UKCP09 grid-squares as spatially coherent with the Ennerdale catchment was examined against monthly (rather than daily) rainfall totals; and the rainfall-runoff transform was rudimentary—excluding land-use factors.

Many of these issues could be resolved with greater investment of time and availability of data and we would note that the simplistic nature of (e.g.) our rainfall-runoff transform function does not detract from our conclusions around the methodology and its benefits. Other limitations have parallels with conventional analysis. The definition of a demand uncertainty distribution, for instance, must always contain an element of subjectivity (Dziegielewski & Bauman, 2011). The wider fundamental question of whether probability distribution functions offer a suitable mechanism for describing system uncertainties undoubtedly calls for further attention. This applies in both supply and demand input parameters. We elaborate on these limitations below in our discussion around scaling up to a risk-based water resources planning standard.

We believe the analysis presented would fall comfortably within the skill-set of today's trained and experienced hydrologist, who typically possesses the statistical proficiency and computer programming abilities needed to set up multiple water resource system simulations and analyse the outputs. The deliberate use of UK industry-standard software, standard water resources planning data, and freely available climate projections shows that the method could be applied—at least in the type of situation in England and Wales illustrated above—with negligible financial investment in data, software or computing power.

3.6.3 Scaling up to a risk-based industry standard

Although the risk-based approach to water resources planning described above is both practically feasible and beneficial, further research is required to demonstrate how it could transcend the Deployable Output approach currently described in the regulatory guidelines. Here we identify several nontrivial challenges that would hinder an industry-wide change in practice.

We previously described the ambiguity around Ennerdale's viability as a supply source and the impending decision of whether to abandon it and mobilise new resources at relatively large cost. Could a risk-based methodology inform this decision? We envisage major problems in using the output from our approach in this way. Our risk profile (refer back to profile c in Figure 3-7) highlights an approximate 5% chance of failing to meet a 1 in 100 year level of service (varying depending on duration below 'hands-off' lake level). The potential consequences of this event include customer supply being cut off, so water companies—and their customers—are likely to seek some level of assurance that failure of this type will be extremely unlikely. They might, for instance, want 95% confidence that the service exceeds a failure return period of 1 in 1000 year. In theory, advanced stochastic models can produce synthetic samples of unlimited length that would enable the analyst to estimate these extremes (albeit with a substantial dose of uncertainty). The UKCP09 weather generator, however, comprises a relatively simple stochastic model calibrated using only 30 years' data and is therefore insufficient for reproducing the extremes that might be more interesting for the water resources planner (e.g. 200-, 500-, 1000-year events). The user is thereby limited to 100-year sequences and the associated guidance cautions against 'extreme statistics for return periods longer than 10 years.' The user-friendly tools for generating stochastic data may therefore be ill-equipped to generate occurrence probabilities for these events (Harris *et al.*, 2013, make a similar point).

We would also face problems when attempting to estimate the value of the consequences. The Ennerdale problem is particularly complex because the worst-case consequence of lake 'emptying' might not occur in reality: the lake depth extends 40m below the 1.7m hands off lake level and, in a situation where the decision weighs drawing the lake down further versus cutting off customer supply, it is hard to envisage supply being cut off. Should a planner then quantify risk based on some planned consequence—which would assume the lake effectively empties at 1.7m depth—or real consequence, which would allow for the more realistic eventuality in which the urgency of the situation overrides the legislative requirement to keep within environmental thresholds? Planning guidelines would need to specify these details clearly.

The allocation of risk poses a further problem. For instance, the risk of the reservoir emptying could fall most heavily on the water supply company (as it would initiate an expensive emergency response) or on the environment (if the company chose to sacrifice either the river flows or the lake-levels) or on the customer (if severe water-use restrictions were imposed). But, in a rational risk-based assessment, these risks should always be weighed against a cost of adaptation that is primarily a risk (in opportunity cost) to the customer—paradoxically the party with both the poorest understanding of the risks and least representation in the planning process. This imbalance of incentives—if, indeed, it exists (and we have little reason to suppose it would not)—could exact a toll on the more 'flexible' elements of the analytical process, such as the definition of a 'tolerable' level or service. Thus we challenge the widely-held assumption (e.g., Lopez *et al.*, 2009) that a rational risk-based analysis naturally leads to a better quality decision: at best it offers a foundation for striving toward that goal.

Hall and Borgomeo (2013) argue that customer research could supply the information required to justify investments for reaching particular service standards. We view this as the most promising and practical way of deploying probabilistic information in a decision framework, but suggest that this approach would open up a range of new challenges for planners. In particular, we question whether a select group of customers would be best placed to define the tolerable occurrence probabilities for failing the most extreme thresholds. Total storage depletion, for instance, would be associated with wider social and economic consequences, such as loss of business functions, impacts on human health, slowing of economic growth (Brown *et al.*, 2013), etc. These impacts lie well beyond obvious inconveniences to individual households. Critically, ignoring the possibility of total storage depletion in the analysis would threaten its transparency: the positions of triggers (e.g., non-essential use ban) affect the probability of running out of water and ought to be seen as both a consequence (customer inconvenience) and a flexible option for alleviating supply failure risk (a non-essential use ban reduces demand, so if it were triggered at a lower storage volume then the probability of full storage depletion would be increased and vice versa). Indeed, a full exploration of risks implies testing for the trade-offs resulting from repositioning triggers (i.e., trade-off risk of restriction against risk of running out of water).

Returning to a technicality, the simple problem in our analysis is represented at a relatively small spatial extent and we foresee several difficulties in up-scaling to larger water resource zones. Firstly, the method requires reasonable rainfall-runoff models for all catchments that feed rivers, reservoirs and groundwater sources in supply systems. This would impose a lengthy and expensive process in terms of human time—although we should note that building these models would, in the main, be a one-off investment and that additional benefits may be realised from the resulting enhanced understanding of catchment response. Secondly, analysis of larger water resource zones would demand greater computer time—a technically difficult but surmountable challenge. In fact, recent studies have demonstrated more computationally-efficient methods for simulating large resource zones to acceptable accuracy (Matrosov *et al.*, 2011) and for using a network of computers to share the burden of multiple simulations (Vamvakeridou-Lyroudia *et al.*, 2009). Lastly, problems may arise in large systems that source water from various sources across a wide spatial extent. To illustrate, United Utilities' Integrated Zone includes reservoir sites more than 200km apart. This presents a technical challenge because the UKCP09 projections cannot be used to generate two sets of stochastic series that link the temporal dependencies between two distant catchments, although we should note that multi-site stochastic models for generating daily rainfall data are available in a relatively user-friendly format (e.g., Srikanthan, 2005). Alternatively, the outputs of the UK 'Future Flows' project (Prudhomme *et al*, 2012) provides spatially linked daily weather projections, although relying on this data would mean trading off against certain features of the UKCP09 outputs that fit better with our risk-based approach (e.g., the availability of up to 10,000 synthetic weather time series).

We previously alluded to issues around the definition of input uncertainty distributions and the resulting impact on the legitimacy of our analysis. These distributions in hydrological and demand input parameters may misrepresent uncertainty, either due to error in distribution parameters or simply because probability distribution functions are, in principle, unsuited to describe uncertainty in these systems. The former of these concerns somewhat discredits the precise probabilities that arise from the analysis. A similar analysis underpinned by different climate models would yield different results (though perhaps not radically so), and so critics might construe our own outputs as potentially misleading. To counter this viewpoint we can claim that the risk-based method at least provides a framework that would easily accommodate new understandings of uncertainty and quickly bring them into the analysis as climate modelling research progresses (although we must concede that a stronger understanding of future demand uncertainty seems beyond reach). This framework embodies the added benefit of clearly laying out the assumptions of uncertainty and, as Hall *et al.* (2012a) suggest, sensitivity analysis of plausible variations in input distributions (perhaps defined using different climate model ensembles) could quantify the relative importance of these assumptions, which might turn out to be trivial.

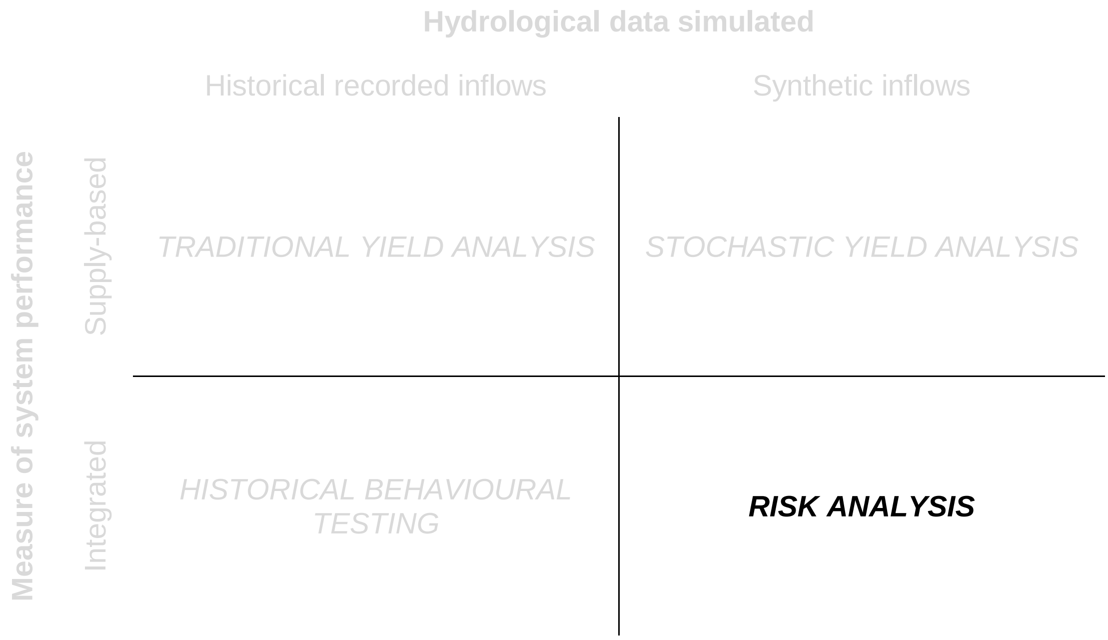
A tougher question is whether uncertainty distributions, in principle, offer a legitimate description of future uncertainty. Highly non-linear systems influence water resources system performance. Climate response, for instance, may shift dramatically in line with currently unrecognised reinforcing system feedbacks that current GCMs fail to represent (Beven, 2011). Probability distribution functions make poor descriptors of uncertainty in systems that throw up surprise events (Taleb, 2007). To some extent, this makes water resources planning a problem of 'deep' uncertainty in which no amount of science can define legitimate limits of input parameters. Brown and Baroang (2011) consider 'surprise' events and suggest possible workarounds to reconcile the resulting dilemma within a risk-based planning approach. These suggestions extend the role of contingency drought planning and emphasise a need for innovative risk management instruments, which might include insurance mechanisms based on weather index derivatives that would pay out on extreme circumstances allowing companies to cut-off and compensate large water users (under some pre-agreement) and/or finance extremely expensive emergency operational measures (e.g., tankering water). We suggest that scaling up to a risk-based industry standard should coincide with a detailed exploration of the feasibility of deploying these instruments, which remain largely unexplored in the context of water resources planning in England and Wales.

3.7 Conclusions

We have aimed to understand whether water resources planning practice in England and Wales would benefit from probabilistic analysis and corresponding risk-based project appraisal for improvements to water resources systems. Our focus on practicalities and industry challenges separates our contribution from similar recent work in this context—we refer readers to Harris *et al.* (2013) for a similar UK-based study that provides more explicit detail around the relative importance of different sources of climate impact uncertainty. We have demonstrated that a risk-based methodology informed by multiple simulations of water resources systems is both practically feasible for a single, simple, water resource zone and beneficial in that it nurtures a more detailed understanding of risks. In this regard it lays the foundations for more rational decision-making. Nonetheless, we identified several non-trivial problems that would hinder implementation of the approach as a water resources planning standard. Some of these limitations are specific to a heavily regulated water service provider, but others apply more widely. We believe further research is needed to demonstrate how these problems might be resolved and how a transition to this approach would transpire. Future studies might aim to: develop practical methodologies for applying these methods to larger resource zones; identify and overcome problems associated with the effective regulation of this methodology, including how to allocate and value risks in economic evaluations and project appraisal; and produce workable methods for dealing with 'surprise' events that would undermine the uncertainty distributions used to inform the analysis.

CHAPTER 4 ROBUST DECISION FRAMEWORKS

Article title	*Challenges for informing water resources planning decisions through "robust decision frameworks"*
Co-authors	Paul Jeffrey (Cranfield University)
Co-author contributions	Corrections and style guidance, general discussion.
Publication status	Under review at *Water Resources Management*. An earlier version was published in the Proceedings of the 8th International Conference of the European Water Resources Association, Porto, June 2013.
Reasoning for case and site selection	The RDM approach had begun to appear in UK-based academic literature and had generated some interest amongst the research sponsors. The study builds on the prior analysis by executing the risk analysis on a larger system with multiple catchment inflows across a wider area.



4.1 Introduction

It is argued widely that water resources systems must be adapted to cope with an uncertain future climate. This is the mantra of the contemporary water resources planning profession and the premise of numerous climate change impact studies. Planners want to know how to reach rational decisions under such uncertainty; consultants and academics want to provide the necessary tools and expertise.

The problem of designing water resource systems for an uncertain climate has provoked a lot of new research from which prospective planning frameworks have begun to emerge. These tend to feature two distinct elements: a stochastic vulnerability assessment—often informed by ensembles of climate model projections—and an accompanying decision analysis that deploys the resulting system performance data to reach so-called "adaptation decisions." The addition of the word "adaptation" presupposes and implies that planners need a new brand of decision when faced with uncertain climate change—an idea that has been contested on the grounds that conventional planning methods have performed adequately amidst the numerous sources of uncertainty and non-stationarity that eclipse human-induced climate change impacts on hydrological systems (Koutsoyiannis, 2011; Lins and Cohn, 2011).

Nonetheless, a debate on whether and how to update existing water resources planning methods has begun to play out in England and Wales, where water service providers must follow a planning guideline prescribed by industry regulators (Environment Agency, 2012). Hall *et al.* (2012a), Hall and Borgomeo (2013) and Borgomeo *et al.* (2014) have described and promoted a "risk-based" planning framework that deploys ensembles of transient climate model data to quantify probabilities of achieving hazard-based service levels through the planning horizon. The approach builds on the ideas advanced by Hashimoto *et al.* (1982a), who proposed metrics of risk (based on stationary stochastic hydrology) that would "supplement other standard project evaluation criteria, including the distribution of project benefits and costs as well as various social and environmental impacts." The feasibility of introducing non-stationarity into the Hashimoto-style risk-metrics has been well demonstrated in the UK context through several studies that have propagated climate model projections through hydrological and water resources system models (e.g., New *et al.*, 2007; Manning *et al.*, 2009; Lopez *et al.*, 2009; Harris *et al.*, 2013), although Turner *et al.* (in press) highlight technical challenges for up-scaling these analyses to deal with the complexities of larger water resources systems, as well as institutional challenges for adopting them as part of regulatory guidelines in a privatised water industry.

The debate has now progressed to focus on how the data generated by exploratory scenario analysis might inform strategy and investment decisions. Hall and Borgomeo (2013) seek to retain elements of conventional planning: tolerable levels of service to be defined through customer research and addressed through interventions that consider risks for various stakeholders. This contrasts with quantitative procedures that aim to integrate the plethora of decision variables in a calculation that returns a favourable solution. These include "robust decision frameworks" that provide a quantitative method for using large sets of system performance data to guide planning and project design under "deep uncertainty" (Lempert *et al.*, 2003; Weaver *et al.*, 2013). These approaches reject the notion of precise probabilities and instead aim to use exploratory modelling—driven by climate models as "scenario generators"—to stress-test the system under a wide range of possibilities. The resulting data are used to select interventions that perform adequately across the range of uncertainties using some form of regret analysis. Whilst definitions of "robustness" vary, the general precept of a robust decision bears similarity to the Hashimoto *et al.* (1982b) definition: ability to cope with a wide range of conditions at little additional cost.

The practice of stress-testing systems through exploratory modelling is nothing new to the water resources planning community; stochastic models for generating multi-site synthetic streamflow traces were developed in the 1950s in order to perform weak-link analyses of large reservoir systems under plausible but unlikely hydrological conditions (Fiering, 1997). Nor does exploratory modelling require climate models to generate scenarios; stochastic models for generating synthetic hydrological data exist in a variety of forms that can encompass—among other complexities—climate non-stationarity (Salas *et al.,* 2012). From the water resources planner's perspective, the novelty of the "robust decision framework" surely arises from the quantitative analytical application of the system performance information, rather than the exploratory analysis that produces it.

Demonstrations of robust decision frameworks in a water resources planning context have begun to emerge in the literature (e.g., Lempert and Groves, 2010; Chen *et al.*, 2013; Matrosov *et al*., 2013a, 2013b). These studies have succeeded in translating theory into case demonstration, revealing the technical feasibility of executing many exploratory scenario simulations on large systems and pulling the results through decision analyses to generate recommendations for policy intervention. However, there remain several weaknesses and unanswered questions. Specifically, given the premise of "deep uncertainty," how can a legitimate range of uncertainty be delimited in the analysis, and how sensitive are the recommendations to those assumptions? Climate model outputs, for instance, represent only the lower bound on the maximum range of climate uncertainty (Stainforth *et al.*, 2007). So when considering robust decision making analyses that rely on these data, one is left pondering whether a different set of "robust" solutions might emerge under the examination of the deeper uncertainties that the methodology sets out to deal with in the first place.

A further criticism considers the "regret" associated with financial cost of different plans. This translates as opportunity cost for bill payers where systems are over-designed, occurring under more favourable future supply and/or demand conditions (e.g., low demand growth; low impact of climate change on drought severity and frequency). Previous studies have dealt with cost using multi-criteria analysis (e.g., Matrosov *et al.*, 2013a); a remaining question is how to quantify suitable weightings. Other studies have confronted this issue by assigning penalty costs (per unit volume) to simulated supply-demand shortfalls (e.g. Pallottino *et al.*, 2005; Chen *et al.*, 2013), although few elaborate on valuation methods for these costs (to which the solution must be highly sensitive) and we find none that consider the prospect of non-linearity of consequence. For instance, a 100% shortfall may be more than twice as damaging as a 50% shortfall. Or perhaps the impact of a week-long supply disruption would be more than seven times worse than the impact of a day-long supply disruption.

This paper seeks to demonstrate where the issues described above begin to erode the utility of "robust decision making" principles. In order to clearly communicate our findings, an analysis that rests on these principles is executed on a simple water resources system. The analysis assumes a small, discrete set of intervention options for addressing supply failure risk is available.

4.2 Test bed—a weakly interlinked water resources system

4.2.1 Model specifications

The test bed for this analysis is a stylised version of a water resources system located in northwest England. The system harvests water primarily from relatively wet catchments (annual rainfall ~ 1800mm) characterised by steep, rocky terrain and corresponding flashy hydrological regime. The system comprises three distinct, weakly-connected supply areas. Water is abstracted from a combination of small storage reservoirs (~90 day critical period under the drought of record), streams and boreholes to supply 55 ML/d (annual average) to around 150,000 people. These water resources are protected under various environmental designations, which severely restrict the range of intervention options available to the incumbent water company. For the purpose of this study, we imposed hypothetical abstraction limits and minimum volumes for compensatory flow releases. This allowed us to replicate conditions similar to those experienced by water companies facing increasingly stringent environmental regulations.

A model of the resource system was prepared using Aquator (Oxford Scientific Software, 2008), which includes the major bulk supply assets as well as various operating rules, such as time limited abstraction licenses, compensation flow arrangements based on reservoir levels, and control rules. The system setup is conceptualised in Figure 4-1. The software simulates the movement of water within the system using an optimiser that minimises costs—computed using marginal costs of use for each component in the model—under normal operating conditions. A breach of any reservoir control curve switches the optimiser mode to maximise "resource state" (e.g. return reservoir levels to a healthy state), which permits the model to draw on more expensive resources to augment storages and serve demands. Specifically, this system drafts water from boreholes and transfers (which augment rather than fully satisfy water demands) when storages begin to draw down during dry conditions. Control curves were positioned based on historical operational practice.

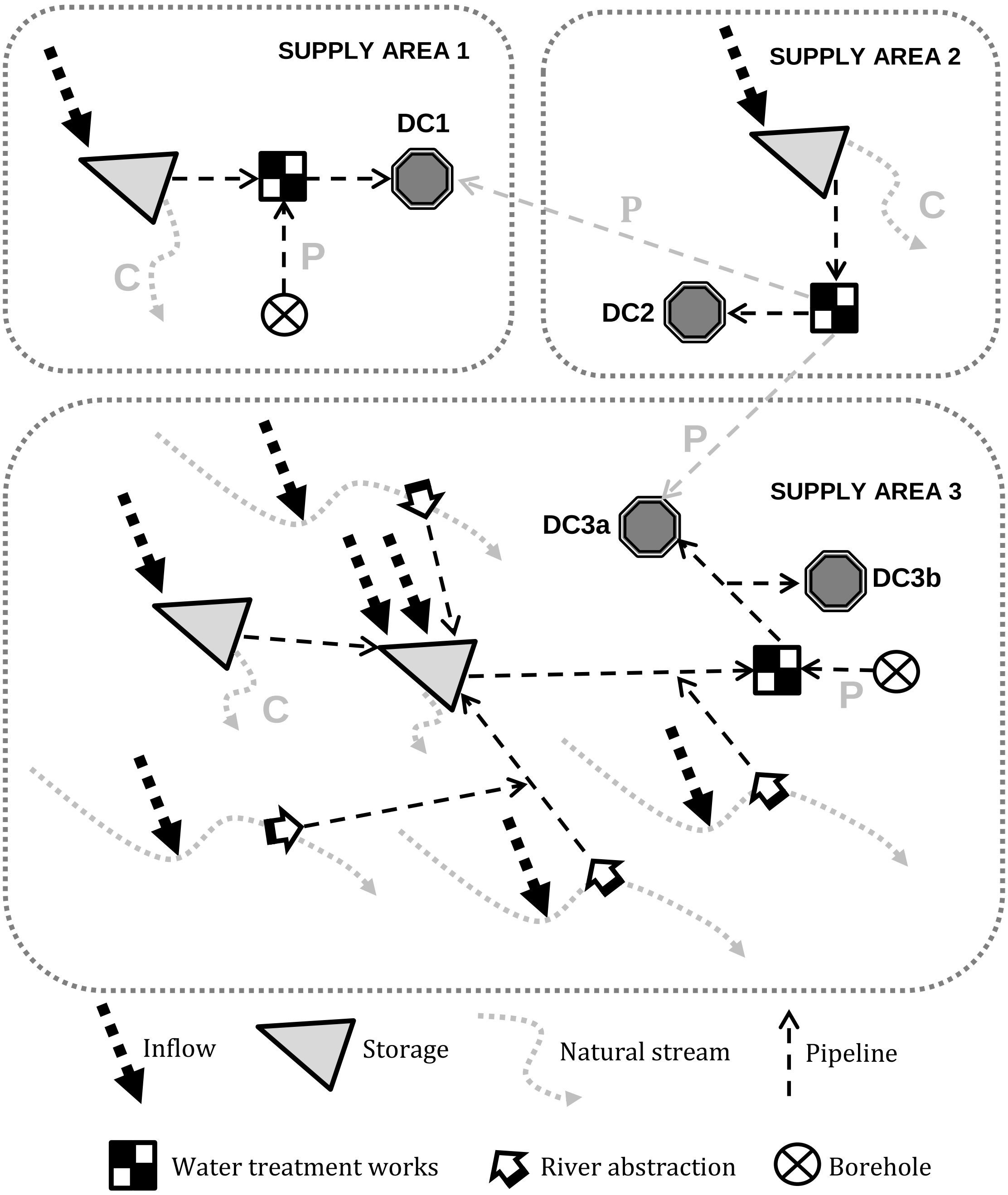


Figure 4-1 Resource system schematic showing reservoirs, inflow sequences (large dark perforated arrows), river reaches, linkages, abstractions, boreholes, treatment works and demand centres (DCs). C denotes compensatory flow requirements on river reaches. P denotes pumped pipelines.

4.2.2 System performance under conventional methodology

Under the existing method of analysis mandated by water resources planning guidelines in England and Wales (Environment Agency, 2012), system performance is measured using a deterministic metric of yield, termed Deployable Output (DO). The DO represents the highest demand that the system can consistently supply under a repeat of the historical inflow conditions, subject to a set of modelled constraints. A reserve storage volume provides a safety buffer to protect against the occurrence of droughts more extreme than those contained within the hydrological record. This form of system yield analysis—and the use of a 'reserve storage'—characterises the planning approach used in practice across much of the developed world (e.g., Erlanger and Neal, 2005; Rush *et al.*, 2011; New Jersey Department of Environmental Protection, 2011; Water Research Foundation, 2013). The planning paradigm is 'least-cost capacity expansion'—meaning the aim is to determine and schedule the least-cost combination of options to address any supply-demand 'deficit' within the planning horizon (25 years in England and Wales). Figure 4-2 shows the system yield for the full test zone and separately for each of the supply areas, based on the baseline system configuration and historical recorded inflows. The yield for the full zone is constrained by the yield at Supply Area 1. This occurs because the system is weakly linked; Supply Area 1 can fail even whilst there is stored water elsewhere in the system. Supply Area 2 is in a state of minor deficit and Supply Area 3 is deemed healthy in a state of 'surplus.' These results provide a baseline against which to test the outputs of our exploratory stress testing analysis.

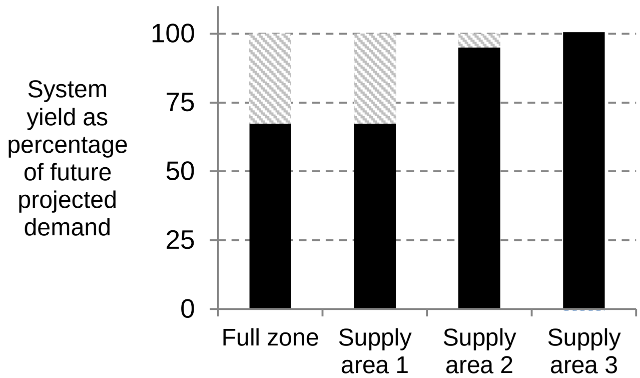


Figure 4-2 System yield for the full zone and separate supply areas.

4.2.3 A discrete set of interventions for addressing supply failure risk

Under the conditions shown in Figure 4-2—i.e., a system yield deficit against projected demand—planners would formulate and model sets of options in an attempt to realign the "supply-demand balance." In this study, the following six options were defined and modelled: (A) Do nothing; (B) A new river abstraction to Supply Area 1; (C) Remote groundwater schemes feeding Supply Area 3, plus a new pipeline to support Supply Area 1 from Supply Area 3; (D) An increase in the transfer capacity between Supply Area 2 and Supply Area 1; (E) A re-opening of an abandoned groundwater source in Supply Area 3 plus new pipeline between Supply Area 3 and Supply Area 1; and (F) A re-zoning of demands in Supply Area 1 so that they are fed from a large neighbouring resource system via inter-basin transfer. The following section describes how a stochastic vulnerability assessment was performed for each option. The subsequent section assesses whether the resulting understanding would endorse a significantly different system design compared to one based on the system yield assessment method described above.

4.3 Exploring the implications of stochastic stress testing on understandings of system performance

4.3.1 Exploratory scenario analysis

In order to incorporate uncertainties relating to inflow variability and climate change, an ensemble of precipitation and evapotranspiration data was extracted from the 'Future Flows Climate' dataset (Prudhomme *et al.*, 2012). These data were derived from a regional climate model (FF-HadRM3-PPE) run under the SRES A1B emission scenario. The dataset features an 11-member daily weather scenario ensemble (1950-2098) scaled to 1km grid squares covering the whole of the United Kingdom. Data corresponding to the catchments in the test system were extracted. The historical sections of these data were sense -checked against rain gauge data for goodness-of-fit with the first and second statistical moments to corroborate with the checks carried out by Prudhomme *et al.* (2012).

Rainfall-runoff models were prepared for each of nine flow series in the test system. Model parameters were optimised using a genetic algorithm in the Rainfall Runoff Library (RRL) modelling software (Perraud *et al.*, 2003); the objective function minimised a Nash Sutcliff metric (Equation 4–1) for which a λ parameter was implemented to concentrate the calibrations on low flows (method in Barma and Varley, 2012). The first two thirds of each record was used for calibration and the latter third for validation of each model. Nash Sutcliff efficiency scores greater than 0.7 were achieved for all calibration and validation periods, which varied between 3 and 30 years in length depending on data availability.

$$ENS = 1 - \{ \Sigma(m_i^{\lambda} - o_i^{\lambda})^2 / \Sigma(o_i^{\lambda} - \hat{o}^{\lambda})^2 \} \qquad \text{Equation 4–1}$$

Where $\lambda = 0.2$, *i* is time reference, *m* is modelled flow and *o* is observed flow ($\hat{o}$ is mean flow form all observations).

The resulting rainfall-runoff functions were used to convert the ensemble of daily precipitation (and monthly PET) time series into an ensemble of stream flow time series. Eleven scenarios were constructed, each 50 years' length representing the time period 2020 – 2069. To clarify, these are transient sequences based on climate model runs rather than stationary stochastic sequences based on a single time slice.

An external procedure, coded in Microsoft Visual Basic for Applications (VBA), was set up to effect control over the water resources system model. The code automated a batch simulation of the system by running each flow sequence five separate times under alternative randomly sampled uncertainties that were pre-defined in probability distribution functions. These included uncertainties around asset constraints, rainfall runoff models (as emphasised by Ajami *et al.*, 2008) and demand forecasts. On termination of each model run, the code extracted daily time series of demand met (as % of projected demand) for each demand centre in the system and pasted in a spread sheet for post-analysis.

To describe and examine the risks, a vulnerability matrix was adopted (Figure 4-3). This would ultimately create a surface in which the z-axis represents probability of occurrence for demand shortfalls from the 2750 simulated years (5 × 11 × 50-year scenarios). Our matrix was populated using an algorithm that searched each year of output series for a demand shortfall, then characterised that shortfall based on magnitude and duration. Since the data are melded into a single matrix, this particular part of our analysis does not recognise the transient nature of the input scenarios. Instead it uses the severe drought events contained within those scenarios to stress-test the system and characterise the corresponding impacts on customer supply.

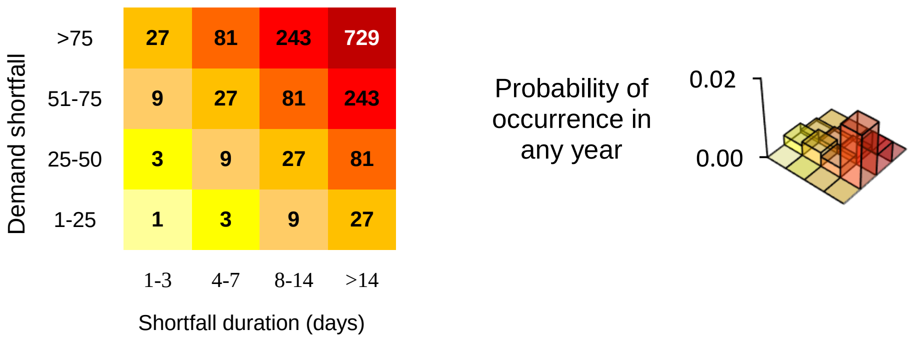


Figure 4-3 Subjective scoring for relative risks and example vulnerability surface based on event probability of occurrence.

In order to understand how this type of risk-based understanding would influence system design, we attached subjective consequence severity scores to each banded square of the vulnerability surface. The scores reflect the assumption that consequence forms a non-linear relationship with both shortfall magnitude and duration. We wish to emphasise at this stage that there exists no recognised methodological approach for integrating the different risk types represented in the vulnerability surface—that is, these subjective consequence scores have no logical or theoretical grounding other than a basic assumption of non-linearity of damage associated with greater shortfall magnitude and duration. (The assumption is that a customer would experience more than twice the inconvenience if suffering a 100% shortfall relative to a 50% shortfall, due loss of basic water needs and so forth). Moreover, there is no question that the results will be influenced by this arbitrary scoring system. The impetus for aim in explicating the vulnerabilities in this way is to highlight a major challenge that is generally glossed over in contemporary studies examining water shortage risk.

The consequence severity scores were used to derive two metrics of risk. The first metric is a relative risk score ('RRS') that quantifies risk in each demand centre (d) by combining probabilities from the stochastic analysis with consequences assigned by the subjective impact scores for each matrix segment (m) and then summing across all segments (Equation 4–2). The second uses these scores to create an overall risk score for each option (j) by combining relative risk scores and weighting them by average annual demand (AAD) for each demand centre in the zone, thereby accounting for population affected by shortfall. We term this the demand-weighted risk score (DWRS, Equation 4–3).

$$RRS_d = \Sigma\ (\text{Probability}_m \times \text{Subjective impact}_m) \qquad \text{Equation 4–2}$$

$$DWRS_j = \Sigma(AAD_d \times RRS_d)\ /\ \text{Total demand} \qquad \text{Equation 4–3}$$

Average incremental costs (£ capital expenditure per unit risk reduced) were computed for each option (based on the demand-weighted risk score) and compared against average incremental costs derived from deterministic system yield assessments (£ capital expenditure per unit increase in DO).

4.3.2 Results from exploratory scenario analysis

Figure 4-4 shows the surfaces derived from the above analysis across the set of intervention options. The graphics identify significant pre-intervention shortfall risks in supply areas 2 and 3—a contrast with the DO analysis depicted in Figure 4-2. Closer inspection of the simulations reveals the reason for these discrepancies: low-probability, high-consequence risks are overlooked by simulations of the short (50-year) historical sequence. In particular, the conventional analysis overlooks a lack of resilience in Supply Area 2, which is vulnerable because it cannot draft water from elsewhere in the system. Also, Demand Centre 3b is vulnerable to significant shortfall in years that contain two separate small droughts, which deplete the annual licensed volume for the borehole in that area. The historical record does not feature these inflow patterns and thus overlooks the corresponding risks.

The granularity of the matrices is interesting because, across 2750 years' simulated inflows, one might expect smoother surfaces with decreasing risk as event duration increases. Instead we find that Demand Centre 2, for instance, exhibits higher risk for long-duration events than short duration events. Of course, one can intuitively reason that short duration events would occur more readily; the result reflects the fact that, despite being 2750 years in length, the inflow sequences contain only a small sample of drought events that cause shortfalls. It so happens that the damaging drought contained within the sequences exceeded 14 day duration. Shortfall magnitude is a different matter: if the reservoir in Supply Area 2 were to fail (by storage depletion) then Demand Centre 2 would immediately suffer 100% shortfall (assuming no new water entering the system) because this population cannot draft water from other sources. Thus, additional inflow data would smooth these profiles along the event duration axis (with decreasing probability as duration increases), but maintain coarseness along the event magnitude axis.

Shortfall-duration vulnerability surfaces* (with relative risk scores)

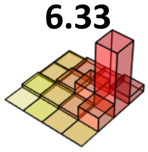
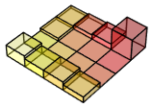
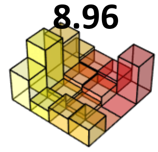
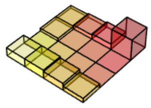
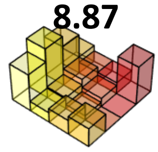
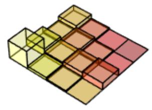
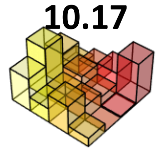
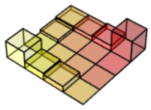
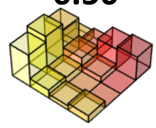
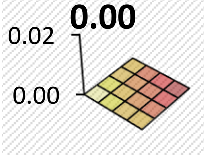
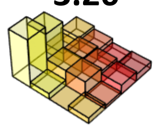
Option	Demand Centre 1	Demand Centre 2	Demand Centre 3a	Demand Centre 3b
(A) 'Do nothing' No investment DWRS = 5.68	**6.33** 	**4.51** 	**0.42** 	**8.96** 
(B) 'River intake' Cost = £65m DWRS = 3.84	**2.71** 	**4.50** 	**0.42** 	**8.87** 
(C) 'Bulk inputs' Cost = £490m DWRS = 1.81	**0.76** 	**4.13** 	**0.00** 	**1.24** 
(D) 'Linkage' Cost = £50m DWRS = 4.57	**3.81** 	**4.55** 	**0.34** 	**10.17** 
(E) 'Link and BH' Cost = £140m DWRS = 4.26	**4.15** 	**4.48** 	**0.03** 	**6.50** 
(F) 'Re-zoning' Cost = £550m DWRS = 0.48	**0.00** 	**0.28** 	**0.02** 	**3.26** 

Figure 4-4 System performance displayed as shortfall-duration vulnerability surfaces [*Prob. scale bottom-left]

Figure 4-5 compares the cost effectiveness of the options against the cost-effectiveness estimates we would derive from the conventional system yield analysis. The “linkage” option is superseded by the “river intake” option as the most cost-effective in the more comprehensive risk assessment. Reference back to the vulnerability surfaces shows how the “linkage” option exacerbates risks in supply areas 2 and 3—a detail that the system yield analysis overlooks. Also, we see the “bulk inputs” option shown as far less cost effective under the risk analysis (it becomes the least cost effective of the options). Again, tracing this back to the vulnerability surfaces we see that this option, despite its cost, fails to deal with the risks in supply area 2.

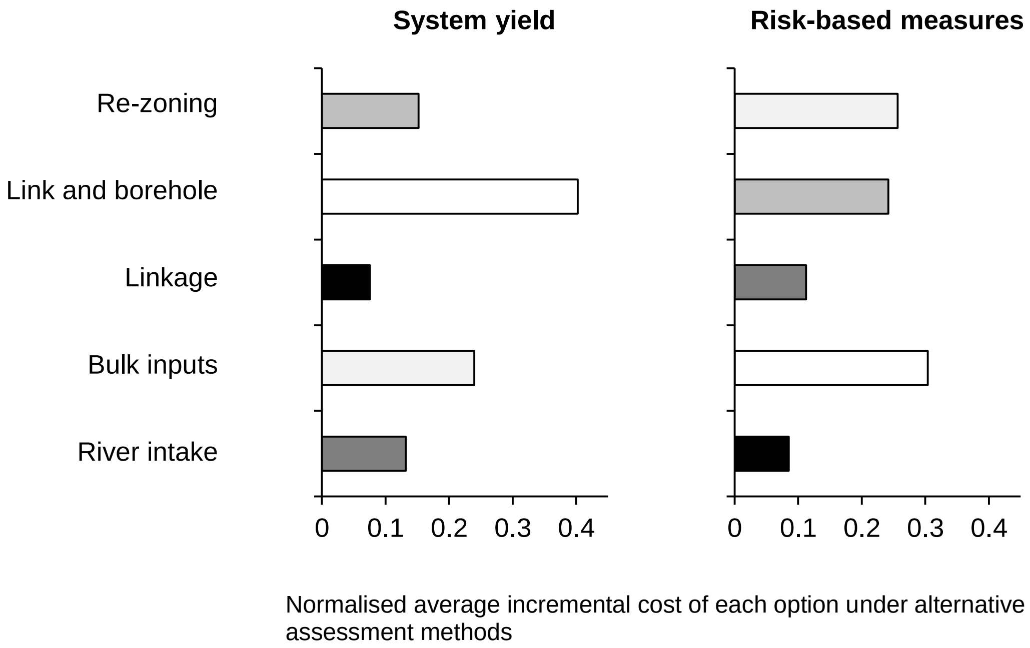


Figure 4-5 Relative cost effectiveness of each option, comparison of alternative assessment methods (shading becomes darker with increasing cost effectiveness).

4.4 Decision analysis

4.4.1 Computing robustness

A quantitative decision analysis was executed using a “minimal regret” criterion (Equation 4–4). The ‘regret’ associated with each option was computed by comparing its performance (denoted RC for ‘Risk Cost’) directly to the other options’ performance under the same scenarios (i.e., same inflow sequence and randomly sampled parameters of demand, asset constraint, etc.).

$$\mathrm{Regret}(j,f) = \mathrm{RC}(j,f) - \min_{i}\{\mathrm{RC}(i,f)\} \qquad \text{Equation 4–4}$$

where *j* is the strategy, *f* is the scenario and *i* indexes through all *j*. Summing across all f provides the total regret for each option. The option with the lowest total regret is considered to be the most robust. The results from the 55 separate scenarios described in the prior analysis were applied in this assessment. To clarify, the risk surfaces shown above were derived from an aggregated understanding of system risk across all scenarios. These data were separated into 55 scenarios to conduct the regret analysis such that the transient nature of the scenarios could be captured through a present value discounting procedure.

The data from the stochastic analysis described above were organised into separate time series of ‘demand met’ (%). We placed penalty values (£) on each risk band and populated the time series with these costs along the full 50-years. A £100 penalty was assigned to each household experiencing a 1 – 3 day demand shortfall of less than 25%. This value was extrapolated up through the risk-bands using the subjective consequence severity scores defined in the above assessment, which translates to an approximate £4.5 billion penalty cost for the full zone for any severe shortfall (>75%) lasting more than 14 days.

The now cost-valued risks were discounted to present values (assuming the start of the time series as present) using a discount rate of 5%, which is a typical value used to discount costs and benefits in engineering project appraisal. This returned a net present risk value and capital investment cost for each option across 55 transient scenarios. These data informed the regret analysis described above, where the total risk cost of each option in any scenario [RC(j,f)] was computed from the sum of the net present value of shortfall risk and the capital cost of the relevant option (capital costs in Figure 4-4).

4.4.2 Results

Figure 4-6 shows the results of the regret analysis compared to probability of being the ‘best’ option. ‘Do nothing’ performs best under approximately 50% of scenarios - in these scenarios the capital investment of the other options is not justified by the level of risk. However, this assessment hides the important point that the ‘do nothing’ option is associated with a high level of regret in some of the scenarios where it fails to perform. Intuitively, these are the scenarios that contain the big drought events causing the high-value risks depicted in our vulnerability surfaces. The regret analysis looks to minimise these regrets and, as such, endorses the ‘river intake’ option. Other options clear up more of the risks, but are associated with greater regrets from opportunity cost in the scenarios where cheaper options would have sufficed.

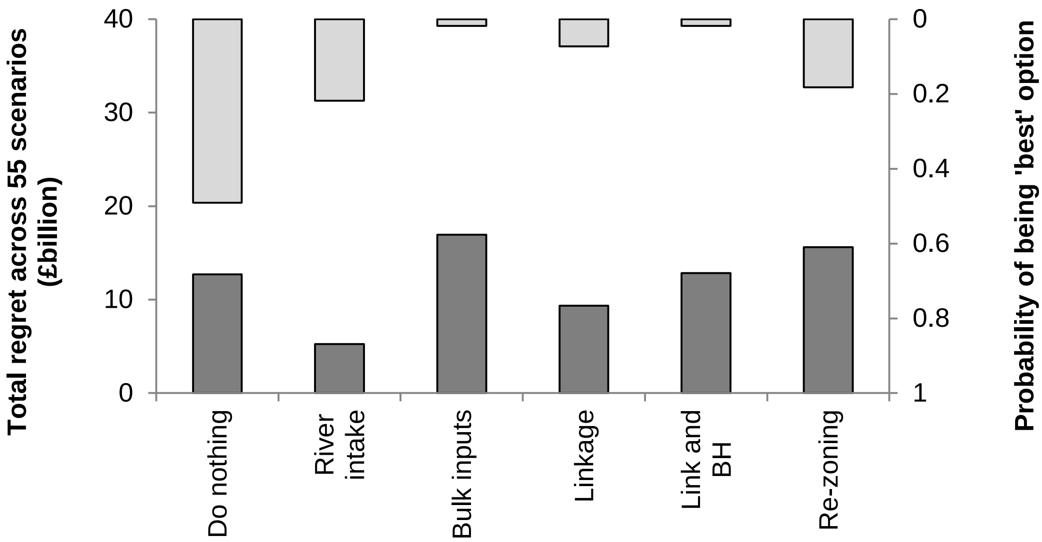


Figure 4-6 Regret (lower, dark) versus probability of being 'best' option (upper, light).

4.5 Discussion—strengths and limitations of a quantitative robust decision making analysis for planning purposes

The results provided in section 4.3 (exploratory stress testing) highlight an important point: the simple process of simulating more streamflow data through the model identifies additional risks that might otherwise be overlooked. We showed how this understanding would influence the planner's understanding of relative cost-effectiveness across our discrete set of intervention options. Critically, we were able to identify which elements of the system created the newly uncovered vulnerabilities. This form of vulnerability analysis would therefore constitute a useful tool for both formulating new options and for comparing their effectiveness for dealing with the supply failure risks. Further analyses of this kind across a wider range of system types could help confirm the generality of this finding.

These results were derived using an 11-member regional climate model ensemble; a multi-model ensemble would no doubt create different risk surface shapes. Perhaps an exploration of more unlikely but plausible extremes would have identified new system vulnerabilities. We highlighted the point that the small sample of severe droughts contained within the streamflow scenarios affected the granularity of the risk surfaces. A simple extension of the model input would create a new picture of risk. One should ask here, then, what is the right input data to describe the uncertainty—and to what extent is this important? It would appear that the key advantage of this analysis is in the identification of previously unrecognised vulnerabilities rather than the quantification of their occurrence frequencies/magnitudes. The key drivers for these vulnerabilities are extreme droughts and unlikely but plausible combinations of inflow variability across the separate catchments contained within the system. Climate models could be useful scenario generators for identifying these vulnerabilities, because they tend to generate more severe and frequent droughts compared with stationary stochastic models. However, even the current state-of-the-art GCMs have failed to adequately reproduce rainfall patterns under historical emissions—particularly for extreme events (Kundzewicz and Stakhiv, 2010; Anagnostopoulos *et al.*, 2010; Beven, 2011). GCMs are therefore an unnecessary—but potentially useful—source of data for stress testing water resources systems.

If we are incapable of accurately defining probabilities then, by definition, we are incapable of quantifying risk in such a way that might justify a given capital spend on system improvements for addressing those risks. Caution against the use of ambiguous probabilities (arising from climate model projections) resembles past concerns regarding the use of failure probabilities computed from stationary stochastic simulations. Those concerns focused on the credibility of the return periods of extreme events that are highly sensitive to arbitrary model parameters. Fiering (1997) noted that stochastic hydrology was developed solely as a means to stress-testing interlinked water resources systems and elaborated on how, if return periods were used to quantitatively inform decisions, a consultant could recommend radically different solutions based on alternative ‘valid’ stochastic series. The benefits of stochastic hydrology were understood to be located in the identification (not quantification) of system risks, which allow planners to form and test intervention options. This line of thinking appears to have been forgotten and perhaps needs reiterating as a new generation of scientists attempt to tackle the same water resources planning problems with ensemble data from climate models.

The accompanying regret analysis calls upon the analyst to value consequences in monetary terms (avoiding this means avoiding the potential ‘regret’ of over-designing, which passes risks onto those who pay for the water supply service in the form of opportunity costs). Various ambiguous assumptions underpin these valuations. This study accounted for the possibility of a non-linear of relationship between consequence value and both shortfall magnitude and duration—an aspect previous studies appear to have overlooked. Yet we concede that there must be incalculable complexities associated with these relationships. Willingness-to-pay surveys could perhaps inform our lower risk bands, but one should not assume that customers could assign credible values to the economic and public health consequences of severely damaging events. The arbitrary assumption of a 5% discount rate compounds these issues (Lind, 1997; Rogers, 1997) and the outputs of the regret analysis would appear to offer little or no value to the planner. The same argument can be made of any form of quantitative risk-based cost-benefit appraisal: the form of analysis will not help overcome weakness intrinsic to the input data, particularly the monetary values placed on alternative consequences.

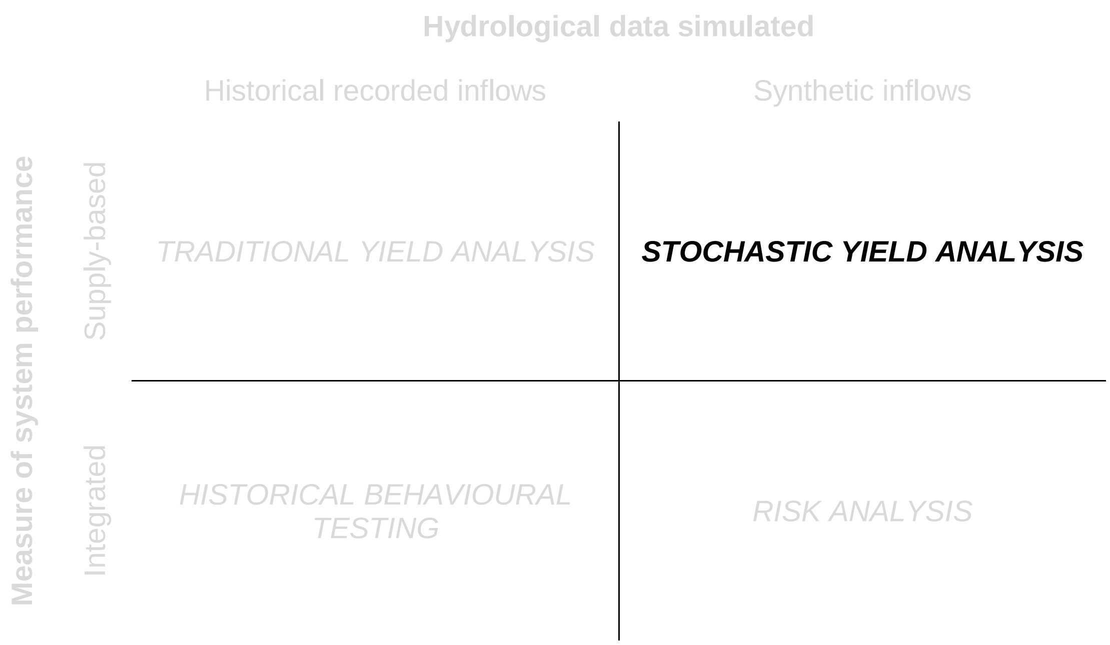
The study showed that the regret analysis recommended a ‘robust’ solution in the sense that the ‘river intake’ option performed adequately across a wider range of scenarios than did ‘do nothing’. The ‘do nothing’ option, in contrast, performs best in most scenarios, but is associated with extremely harmful consequences in a small number of scenarios. The regret analysis appears to have been effective in ruling out a potentially damaging course of action. But would a planner need a regret analysis to reach that decision? The results reported in this study suggest not. The conventional yield analysis (Figure 4-2) rules out “do nothing” for the simple reason that a repeat of the worst historical drought would cause a breach the emergency storage under this plan. Similarly, the more nuanced assessment of risk identifies several high-consequence hazards that we suspect water resources planners would be uncomfortable accepting and so ‘do nothing’ would be rejected on that basis. So, in this case, there appears to be little or no additional value in recommendations provided by the regret analysis.

4.6 Conclusions

We have executed a GCM-informed stochastic vulnerability analysis and accompanying cost-benefit appraisal using a criterion of minimal regret. The stochastic vulnerability analysis highlighted new system risks that an analyst might overlook if using a more conventional deterministic assessment. We showed that this type of assessment would drive new understanding around which infrastructural options perform best in terms of £ per unit reduced risk in the particular system under investigation. However, the quantitative regret analysis relied on various arbitrary assumptions, which would appear to invalidate its claim as a more rational method for justifying system investment compared to the conventional method, which defines a lower-boundary reliability condition that system capacity must satisfy at least cost. We therefore conclude that the benefits of a robust decision making framework are confined to identifying and stress-testing alternative options for system improvement—but do not stretch to justifying investments through regret analyses. Climate model projections may offer little additional value in these water resources planning frameworks because stochastic stress-testing (supplied with scenarios defined using stationary or non-stationary stochastic models rather than climate models) could have identified similar risks to those found in the vulnerability assessment.

CHAPTER 5 DECISION SCALING

Article title	*Linking climate projections to performance: a yield-based decision scaling assessment of a large urban water resources system*
Co-authors	David Marlow (CSIRO), Marie Ekström (CSIRO), Bruce Rhode (Melbourne Water), Udaya Kularathna (Melbourne Water), Paul Jeffrey (Cranfield University)
Co-author contributions	Data provision, style guidance, general discussion. Marie Ekström contributed two paragraphs on climate models.
Publication status	Published in *Water Resources Research*
Reasoning for case and site selection	The decision-scaling approach had received very little attention in the water industry in England and Wales and appeared to be generating interest elsewhere. The use of a large conjunctive use system—located in Australia—extends the range of the thesis and conclusions.



5.1 Introduction

Urban water service providers are responsible for the secure and efficient supply of water to households, businesses, public premises and industry. This requires careful planning and management of bulk supply systems that source, store and move water in such a way that ensures continuity of supply despite highly variable climate conditions. These systems mitigate the substantial societal risks associated with prolonged periods of climatological drought and coinciding water shortage.

Improving water resources systems in an efficient manner—avoiding costly over-design—requires foresight around the likelihood and possible future timing of major droughts. Unfortunately for planners, the capricious nature of extreme climatic events renders the design problem intractable. Hydrologists have long acknowledged the liabilities of event frequency estimation techniques and so water resources systems have been sized up with redundancies to satisfy the planner's tendency toward robust design (Matalas and Fiering, 1977; Salas, 2013). The prospect of climate change exacerbates this issue because the impacts of planetary warming on local hydrological systems are highly uncertain. Rainfall, for instance, can respond to climate change through shifts in the surface and tropospheric moisture and energy budgets (Allen and Ingram, 2002; Held and Soden, 2006), prevailing synoptic circulation systems (e.g., Giorgi and Lionelle, 2008), and various other complex mechanisms. Much of the contemporary literature in this field labels climate change uncertainty as 'deep', 'severe', 'Knightian', etc. (e.g., Adamson *et al.*, 2009; Lempert and Groves, 2010; Hall *et al.*, 2012b; Dessai *et al.*, 2013), reflecting the inadequacies of model-based understandings of complex climate systems (Stainforth *et al.*, 2007; Daron and Stainforth, 2013).

Paradoxically, academic research has focused on GCM-led quantification and characterization of climate change impacts; relatively few water resources studies have looked to develop risk-based methods for dealing with uncertain climate information (Salas *et al.*, 2012). A handful of recent suggestions include real options analysis (Borison *et al.*, 2008), robust decision making (RDM) (Lempert and Groves, 2010), and info-gap theory (Ben-Haim, 2006), although a clear leader among these ideas has yet to surface. Instead, there is a view that the best approach is impossible to identify and that planners ought to select from the range of methods depending on circumstance (Waage and Kaatz, 2011; Barsugli *et al.*, 2012; Hallegatte *et al.*, 2012).

Research into the approaches listed above has continued and case study attempts are beginning to appear in the literature (Jeuland and Whittington, 2014; Matrosov *et al.*, 2013a, 2013b; Korteling *et al.*, 2013). Further method development and case-demonstration may ultimately yield an approach that generates planning decisions that are demonstrably more robust than those generated using conventional planning methods that apply safety margins to buffer against uncertainty. But we should note that the conventional methods have historically served society well against a backdrop of unknowns and uncertainties, including hydrological nonstationarity caused by bushfires, land-use change, complex oceanic-atmospheric factors, and so on (Stakhiv, 2011; Salas *et al.*, 2012). It has been argued that the spectre of human-induced climate change merely extends the existing problem and that its presence warrants little more than additional humility in setting the safety margins (Lins and Cohn, 2011; Matalas, 2012). This conservative view need not disregard the type of stress-testing exploratory analysis espoused by the climate change adaptation community (e.g., Dessai *et al.*, 2009; Weaver *et al.*, 2013). On the contrary, stochastic models for generating synthetic streamflow replicates were developed for that exact purpose (Maass, 1962; Fiering, 1997). These models now exist in a plethora of stationary and non-stationary forms that enable planners to thoroughly assess system vulnerability within the conventional mode of planning. For instance, a safe yield metric based on the 100-year critical drought can be re-examined using hundreds of synthetic traces to produce a distribution from which some pre-agreed percentile and corresponding drought forms the basis for design. In this way, the vulnerability assessment strengthens the design.

Yet, if perceived climate change risk is to be tempered with ‘humility’ (i.e., an extension of system redundancy) then improvements to system design ought to be reasoned using some assessment of the potential impacts. Indeed, Lins and Cohn (2011) acknowledge that ‘one might want to recognise the increased uncertainty that potential anthropogenic influences on climate change introduce into the analysis.’ The question then is how best to estimate that uncertainty and introduce it into the existing planning philosophy. ‘Top-down’ methodologies have dominated the climate change adaptation literature over the last decade (Brown *et al.* 2012). These assessments begin with the climate models and propagate projections through hydrological then water resource system models to assess climate change impacts on system performance. Contradicting this trend, Vogel (2001) and McMahon (2007, 2011) linked reservoir reliability to climate conditions in regional studies, providing a grounding for climate scenario assessment independent of complex procedures for introducing climate change influences into streamflow time series (e.g., weather generators plus rainfall runoff modelling, perturbing streamflow sequences, etc.).

Decision scaling applies these ‘bottom-up’ principles to specific water resources systems to establish decision-critical thresholds in terms of the corresponding climate conditions. Built on the premise that GCM-based climate projections offer a constricted view of possible future climate states (Stainforth *et al.*, 2007), the decision scaling approach avoids using these data to identify vulnerabilities. Instead, the analyst conducts a more extensive vulnerability analysis to understand the sensitivity of the system to changes in climate variables. A climate response function is then constructed to define system performance—and relevant planning thresholds—in terms of the underlying climate statistics that would cause those conditions to materialise. By describing the system performance thresholds in climatic units, the process empowers the planner to assess climate risk by drawing in and comparing various sources of climate information. This may include, but is not limited to, GCM-based projections. It has been argued that such an approach would enable planners to either discount the need for GCMs in cases where the system performance is found to be poorly linked to climate variables, or tailor specific decision-critical thresholds to fit with the available climate information (Brown and Baroang, 2011; Brown and Wilby, 2012).

Recent case studies have highlighted the feasibility of decision scaling in a variety of circumstances (Brown *et al.*, 2011, 2012; Hallegatte *et al.*, 2012; Moody and Brown, 2013; Ghile *et al.*, 2014), although only one of these focuses on a municipal supply system and, in that case, the authors acknowledged a need for further demonstration in 'large and complex systems with multiple performance metrics', which 'often involve greater complexity for the decision maker and for the modelling of the system' (Brown *et al.*, 2012). In this paper, we aim to provide such a demonstration through the development and application of a decision scaling method applied to a large bulk supply system. We then consider the merits of our version of this approach as well as the potential challenges for wider application in urban water resources planning problems.

5.2 Application to the Melbourne bulk water supply system

5.2.1 The Melbourne water supply system

The Melbourne bulk water supply system serves three retail water companies in Melbourne that supply over four million people across a large metropolitan area (Figure 5-1). Smaller volumes of water are also apportioned to regional water authorities, environmental purposes, and irrigation. Water is harvested primarily from protected catchments in the Thomson and Yarra River Basins east of the city, although the system can also source water from a seawater desalination plant, and during critical conditions via inter-basin transfer from Goulburn River. The system storage capacity is approximately 1812 GL, equal to approximately five years' demand. The ten surface water storages include the Thomson reservoir, which has a capacity of approximately 1068 GL (~60 % total system storage).

The Melbourne system is particularly suitable for this study because it features various aspects typical of more complex urban water resources systems. These include: a mixture of climate-dependent and climate-independent water resources; widely dispersed source catchments (e.g., >100 km separates the catchments feeding the Thomson and Yan Yean reservoirs); large inter-annual carryover storage capacity; various environmental flow requirements in rivers/streams affected by the system; a large network of gravity and pumped aqueducts and pipelines connecting the major sources and storages; and demand spread across a wide geographic area.

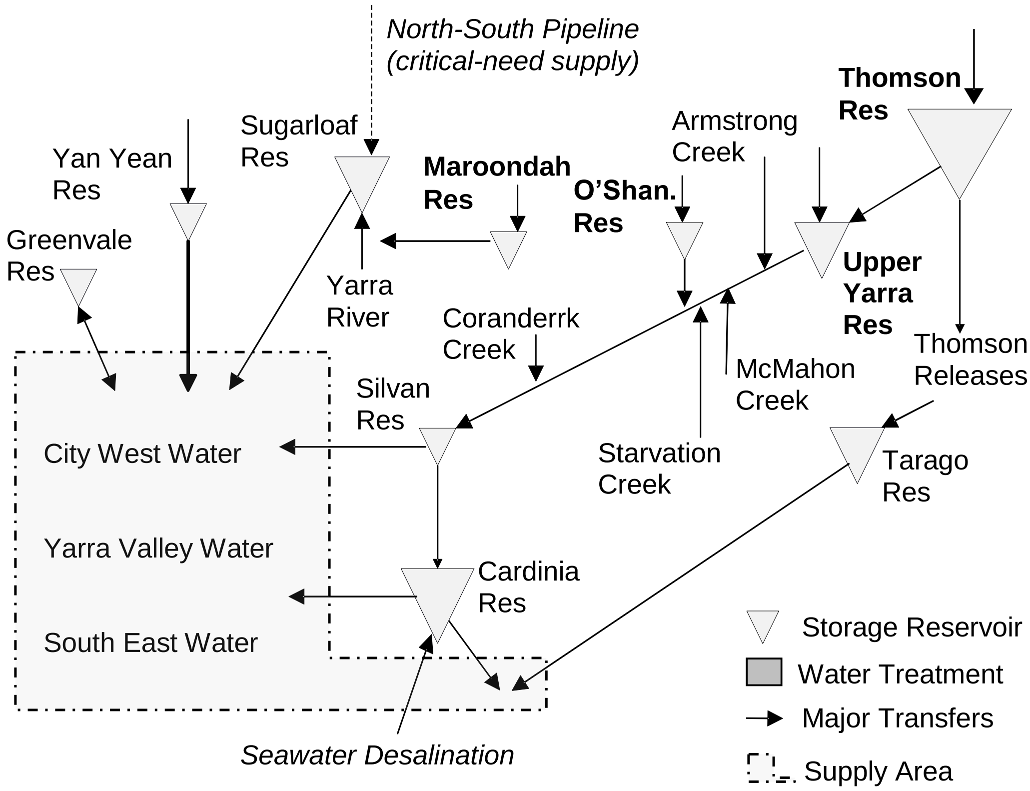


Figure 5-1 Schematic of the Melbourne bulk supply system. Storages in the four major catchment area are labelled with bold type.

5.2.2 Modelling platform and data

The system specifications were incorporated into a model of the supply system using eWater Source – a state-of-the-art node-link mass-balance simulation software package (Kelley and O'Brien, 2012). Physical and operational parameters were written into the various model components. These included reservoir volumes and control curves, pipeline/aqueduct maximum flow constraints, downstream minimum flow constraints, and pump capacities. The model was set up to read in eleven inflow sequences to feed the reservoirs and river reaches represented in the system. Seasonal demand patterns were assigned to each of nine demand nodes as factors that would be applied to the average annual demand. The software applies a linear network solver (RELAX IV – Bertsekas and Tseng, 1994) to resolve the movement of water within the system at each time step, where relative penalty costs assigned to different components guide water transfers to replicate the way operators would move water around the system.

Monthly streamflow records were available for all eleven inflow sites for 1913 – 2012 (100 years). Measured annual precipitation totals (mm) and annual averages of daily mean temperatures (ºC) were available as spatially averaged data for the area of four major sub-catchments supplying the system. These sub-catchments, which contain four reservoirs providing more than two thirds of total system storage capacity, are located in Thomson and Yarra River Basin and account for the majority of supply to the system. Patterns of regional change as projected by models in the third Coupled Model Intercomparison Project archive (CMIP3) (Meehl *et al.*, 2007) were extracted from the OzClim platform (Ricketts and Page, 2007) for a polygon area overlying the major sub-catchments. Changes in mean annual temperature (MAT - ºC) and mean annual precipitation (MAP - mm) were extracted for four 30-year time periods centred on 2025, 2035, 2045 and 2055 relative to the baseline period 1974 to 2004. These data comprised 138 projections, based on 23 climate models run under six emissions scenarios. The emissions scenarios used here are the A1B, A2, B1, B2, A1F1, and AT scenarios detailed in the Special Report on Emissions Scenarios (SRES) of the Intergovernmental Panel on Climate Change (Nakićenović *et al.,* 2000). Lastly, a thirty-year demand forecast was obtained, accounting for trends in demographics and housing, industrial and commercial uses, water conservation technology uptake and leakage.

5.2.3 Defining the decision threshold

The decision scaling approach begins with a definition of a vulnerability threshold based on a measure of system performance. Previous applications of the method have used 'reliability' (Brown *et al.*, 2012), but it is often the case that water service providers use multiple criteria to constrain a yield metric of system performance. Yield here is defined as 'the average annual volume that can be supplied by a water supply system subject to an adopted set of operational rules and a typical demand pattern without violating a given level of service standard' (Erlanger and Neal, 2005). The decision threshold—the point at which the planner would consider intervening—occurs where yield is exceeded by demand, opening up a potential for a 'deficit' between supply and demand that may manifest during drought periods.

In this study, yield was constrained using two service-based criteria: a minimum reliability criterion (0.95) based on a water-use intervention threshold as detailed in the Water Outlook published by the Melbourne water companies on 1 December each year, and a vulnerability criterion based on a maximum allowable drawdown in total system storage. Both trigger points were based on a total system storage (TSS) time series output from a 100-year simulation. Figure 5-2 illustrates two separate TSS outputs from 100-year simulations, highlighting how failure could occur in both cases. Note that violation of the intervention-based criterion could arise from consecutive years below the intervention trigger or separate droughts. It should also be noted that Melbourne Water currently uses three separate service criteria to define yield, and so the results reported in this paper are not directly comparable to other studies of climate impact on the Melbourne system.

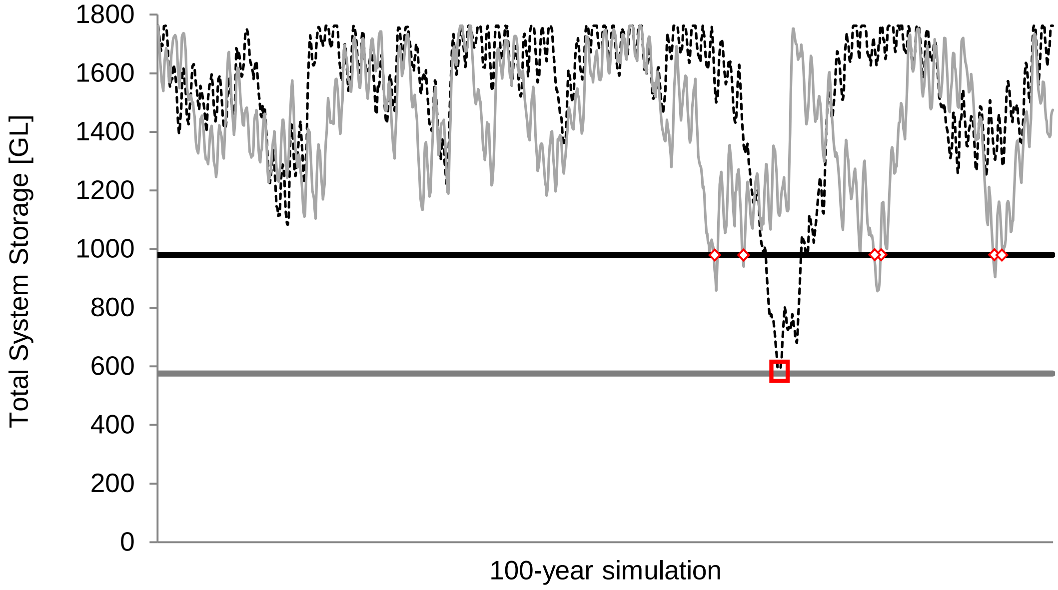


Figure 5-2 Two alternative TSS output time series, each failing under a separate criterion. The lightly shaded series fails under the 'reliability' criterion (> five separate years below 980 GL). The dashed series fails under the 'vulnerability' criterion (TSS falls below 575 GL - in the example the lower trigger is breached within five years of crossing the upper trigger, otherwise it would have failed under the reliability criterion before hitting the vulnerability threshold).

5.2.4 Creating a climate response function

Using the eleven-site 100-year historic inflows, a 100,000-year annual multi-site synthetic streamflow trace was generated using a multivariate autoregressive (lag-1) model (Sveinsson *et al.*, 2007). Lane's method of fragments was applied to temporally disaggregate the data set to a monthly time step (Lane, 1979). The multi-site sequence was then divided into 1000 separate 100-year traces for which yield would be computed and compared to the underlying flow statistics.

Before proceeding with this model, checks were made to ensure it reproduced droughts of a similar severity to those contained within the historic sequence. In particular, an adequate climate response function would require stress-testing under severe multiyear droughts comparable to the so-called 'Millennium drought', which has been described as the worst recorded drought in southeast Australia (van Dijk *et al.*, 2013) and is represented in the latter years of the historic sequences used in this study. Drought statistics were derived from each of the eleven annual replicate sequences and compared against the corresponding statistics derived from the historic sequence. The statistics included: the maximum run-length, in years, during which annual inflow is exceeded by the sample mean annual inflow; the maximum deficit volume experienced in a run of consecutive years during which inflow is exceeded by mean annual inflow; and the required Rippl no-failure storage—computed with the sequent peak algorithm—to meet a draft equal to the mean annual inflow (Sveinsson *et al.*, 2007).

Across the eleven inflow sequences, the mean percentage of replicates exhibiting more extreme drought statistics than those of the historic sequences were 39.5% (st. dev = 25.4), 25.4% (st. dev. = 19.0) and 38.2% (st dev. = 11.5) for the maximum run-length, maximum deficit volume and no-failure storage statistics respectively. It was clear on this evidence that many of the sequences did contain the extreme drought characteristics necessary for adequate stress-testing of the system.

An algorithm was coded in the R statistical programming language and environment (R Core Team, 2014) to call and control a simplified version of Melbourne Water's bulk supply system model. The algorithm computed yield for each inflow sequence set using the bisection method to iteratively converge on the highest annual average volume of water the system could supply continuously without violating the two yield-constraining criteria outlined above. For each replicate modelled, the code captured the resulting yield, the output TSS time series of the final yield simulation run, and various statistics of the input flow trace. Specifically, annual and monthly mean, standard deviation, skew, and lag-1 autocorrelation were captured for the net flows entering the system via the four major sub-catchments.

Figure 5-3a shows system yield plotted against the mean annual flow of the four major catchments (MAF-4). Mean annual flow would be expected to be the primary driver of yield given the over-year storage in the system. A stepwise regression eliminated the seasonal statistics (e.g., mean of January flows, etc.) to isolate the remaining influential parameters, which were the standard deviation, skew and lag-1 autocorrelation of the annual flows (see Table 5.1). Using the results of that regression, the impacts of these statistics were accounted for and set to historic values (similar to Brown *et al.*, 2012), leaving a much improved relation between MAF-4 and yield (Figure 5-3b).

Table 5.1 Regression results for MAF-4 parameters relevant to system yield

Parameter	**Coefficient**	**Strd Error**	**T Stat**	***p*-Value**
Intercept	183.71	12.11	15.17	6.37e-47
Mean	0.001145	2.40e-05	47.75	1.24e-259
Standard Dev.	-0.000973	4.97e-05	-19.55	2.72e-72
Lag-1 Autocorrelation	-91.56	6.44	-14.22	6.25e-42
Skew	19.03	2.44	7.82	1.40e-14
				[Adjusted R-squared = 0.7388]

The linear relation suggests that a given yield could be read off the x-axis to predict—with reasonable accuracy ($R^2 = 0.74$)—the mean annual flow conditions in the four major catchments that would cause that level of performance. Equation 5–1 shows the resulting linear relation (Figure 5-3b).

$$\mu' = 0.65\,D + 158.52 \qquad \text{Equation 5–1}$$

where

μ' = the mean annual flow statistic (GL/yr), based on the net inflows entering the system via the four major catchments, and D = system yield (GL/yr).

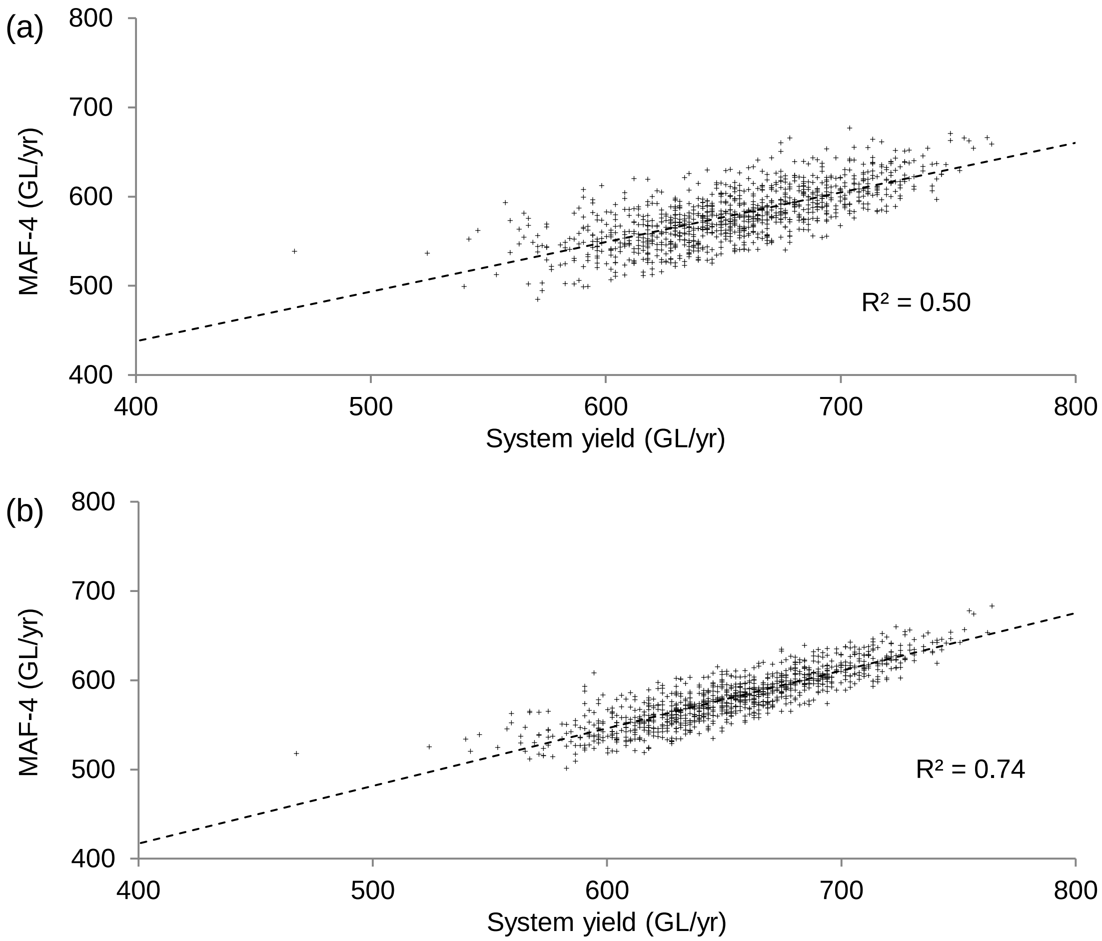


Figure 5-3 Mean annual flow (MAF) – based on the net of the flows entering the system via four major catchments – and resulting yield from 1000 yield search simulations of the bulk supply system. Graph (a) is based on the mean of the annual flows; Graph (b) is based on the same data, but with the other relevant statistics of the annual series (standard deviation, skew, lag-1 autocorrelation) set to historic values using a stepwise regression.

Figure 5-4 shows how the yield defining criteria cause noise in this relation. The graphs show that a given yield could be caused by a range of mean annual flow conditions depending on the characteristics of the drought(s) that breach the restriction threshold. Indeed, the historic observed sequence (MAF-4 = 580 GL/yr) generates a yield of approximately 575 GL/yr, which lies slightly above the regression line. But contrary to eroding the credibility of analysis presented here, this issue exposes an important discrepancy that would arise if yield were assessed on the basis of the historical record alone: if the historic record happened to be an outlier (i.e., reliability driven by a single drought with more than five consecutive years below the threshold or, conversely, more than five distinct droughts each breaching the threshold for less than a year) the estimate of yield would be skewed and unrepresentative of the actual possible flow conditions that the climate could plausibly generate. The relation derived in this analysis instead finds some mid-point between the possibilities and clearly highlights the implications of using the resulting relation in the remainder of the analysis.

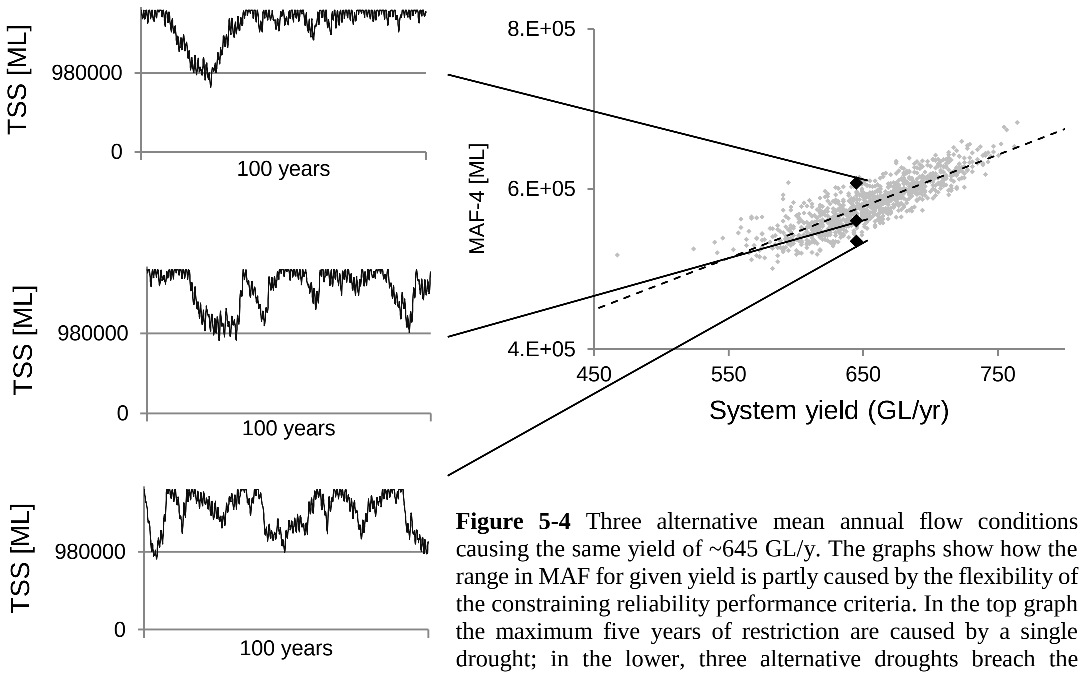


Figure 5-4 Three alternative mean annual flow conditions causing the same yield of ~645 GL/y. The graphs show how the range in MAF for given yield is partly caused by the flexibility of the constraining reliability performance criteria. In the top graph the maximum five years of restriction are caused by a single drought; in the lower, three alternative droughts breach the threshold.

The relation between MAF and yield enabled the creation of a function linking climate conditions to system performance. This is the ‘climate response function’. Developing it required a relationship between the annual flow and the annual climate statistics of the Yarra Ranges. Fifty years of recorded streamflow, precipitation and temperature data for the area of the four major catchments were used to derive the following empirical relation using multiple log-linear regression:

$$\mu = 20.58 \times P^{1.89} \times T^{-1.49} \qquad \text{Equation 5–2}$$

where

μ = the annual flow (ML/yr); P = the spatially averaged annual precipitation totals (mm); and T = the spatially averaged annual average of daily mean temperature (ºC). Figure 5-5 displays the empirical model.

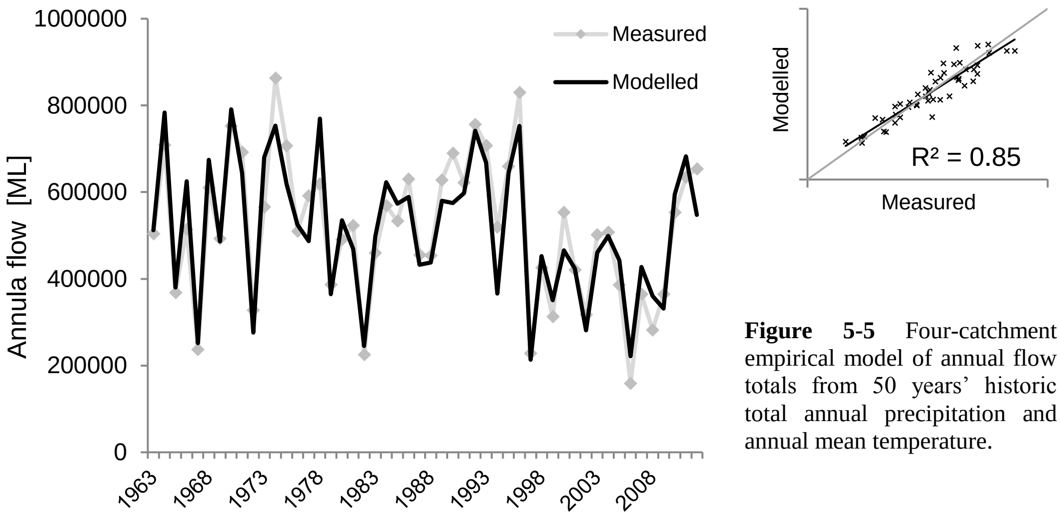


Figure 5-5 Four-catchment empirical model of annual flow totals from 50 years' historic total annual precipitation and annual mean temperature.

Using a normal distribution, with mean and standard deviation of the historic data, a large set of 50-year replicate samples of P and T were generated and used to determine whether Equation 5–2 would hold for the means of these parameters over a long period. This process identified the need for a slight bias correction, which is implemented in Equation 5–3 along with a unit conversion factor for ML to GL.

$$\mu' = 1.024 \times (20.58 \times MAP^{1.89} \times MAT^{-1.49}) \times 10^{-3} \qquad \text{Equation 5–3}$$

where

MAP and MAT are the mean annual precipitation (mm) and mean annual temperature (ºC) respectively.

Combining Equation 5–1 and Equation 5–3 gives:

$$D = (0.033 \times MAP^{1.89} \times MAT^{-1.49}) - 245.38 \qquad \text{Equation 5–4}$$

This is then the climate response function for this particular water resources system, presented as a surface in Figure 5-6.

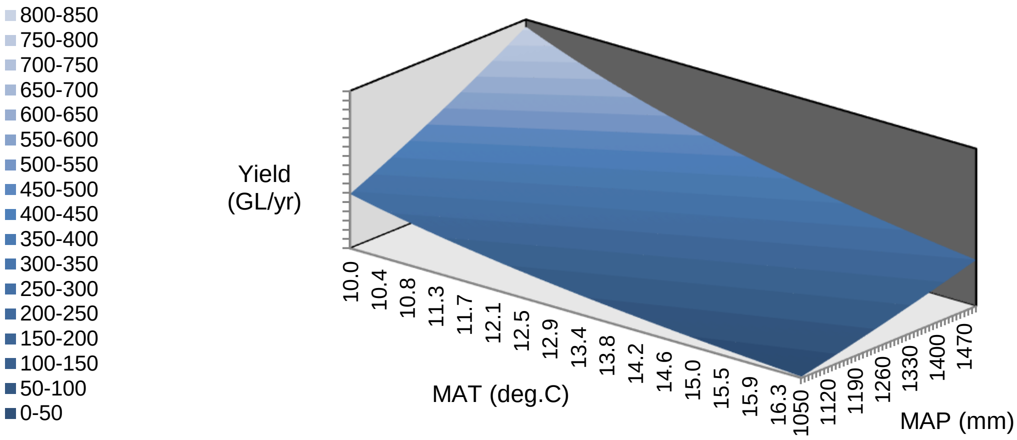


Figure 5-6 Climate response function. Yield is described in terms of the climate conditions that determine its magnitude.

5.2.5 Characterising climate risk

The climate response function in Figure 5-6 was transformed such that it relates yield to change in MAT and MAP compared to a baseline climate. This step was necessary to align the format of the climate response function with the available climate projection data described above. Thus, MAT was expressed in terms of absolute change in ºC and MAP in terms of percentage change, both relative to the corresponding statistics derived from the 30-year time slice centred on 1990. To allow time for planning and implementation, interventions would be generally required at or before the point where yield is exceeded by demand. Therefore, yield can be equated to the forecasted demand and the climate response function can be used to determine combinations of MAT and MAP changes that would cause concern and action at specific points in future. This is a particularly useful way of describing the decision threshold because it enables the analyst to simply plot a climate projection—in the format produced by the GCMs (regionally scaled)—to understand system vulnerability should that projection materialise.

5.3 Results

5.3.1 Climate impact assessment

Figure 5-7 shows the graphical output of the analysis under all four future time periods assessed. Decision thresholds are represented by the grey downward sloping lines; they imply that climate changes represented to the lower left of the threshold could be tolerated (i.e., required levels of service would be exceeded) and climate changes on the upper right would be intolerable, bringing service levels below required standards. Each plotted data point represents a climate model projection. Demand forecast uncertainty was assumed to expand out from 0 % in 2015 to 10 % either side of the forecast by 2055. Two separate decision thresholds are provided in each plot to demarcate the upper and lower bounds of this demand forecast uncertainty. The resulting gap between the decision thresholds expands out toward the lower-left corner of the plot as forecasted demand magnitude and uncertainty increases through the planning horizon. This trajectory reflects the basic principle that a higher demand reduces the robustness of the system to climate change.

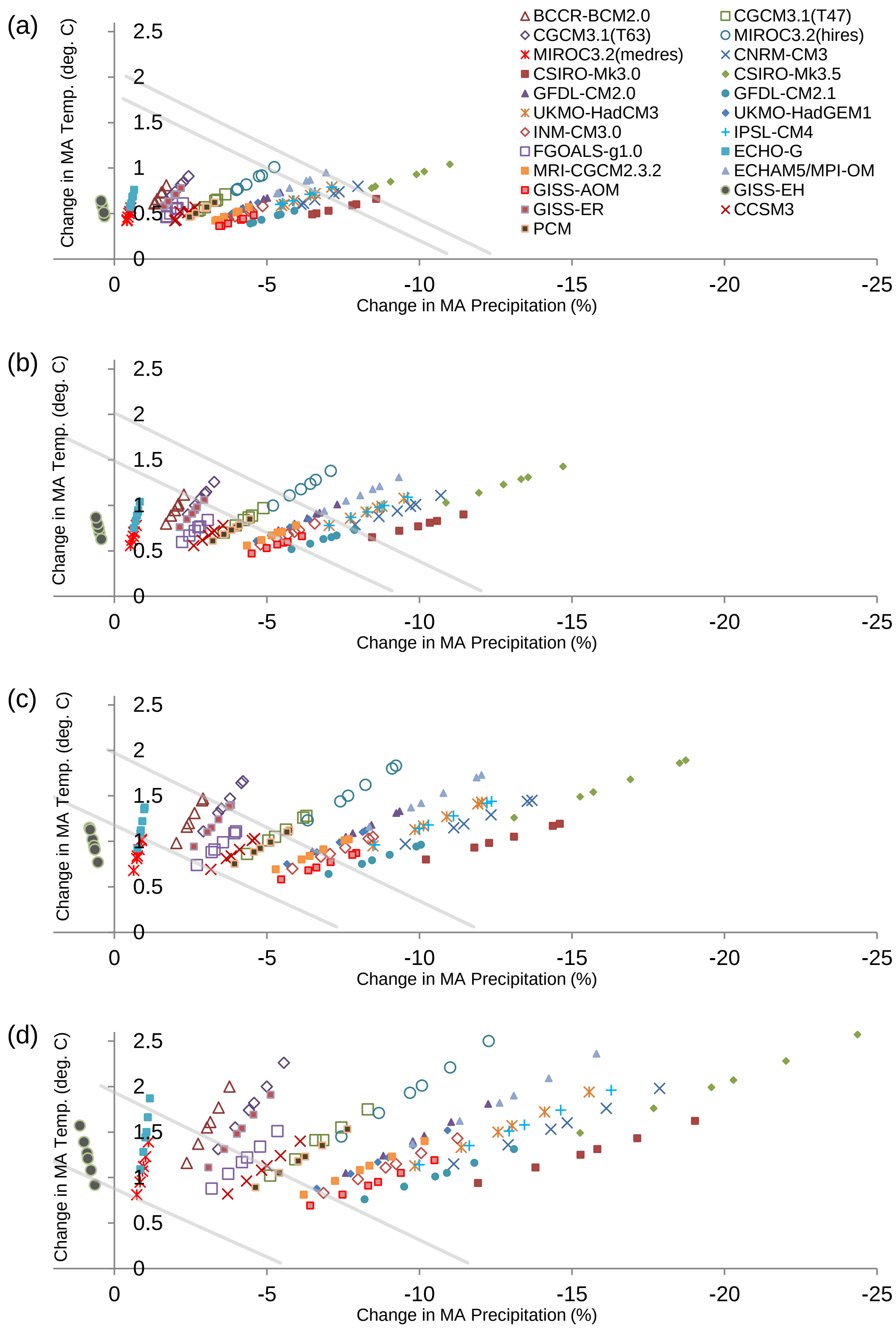


Figure 5-7 Climate projections from 23 GCMs (and six SRES scenarios) plotted against the decision threshold (grey lines marking upper and lower demand forecast scenarios). (a) 2025; (b) 2035; (c) 2045; (d) 2055.

Figure 5-8 summarises the data displayed. We use the terms ‘robust’, ‘unclear’ and ‘insecure’ to denote the position of a climate projection relative to the decision threshold. ‘Robust’—lying lower-left of the threshold—implies that the system can uphold service standards without augmentation; ‘insecure’—lying upper-right of the threshold—implies the opposite. ‘Unclear’ denotes those projections lying between the upper and lower thresholds: the projections could imply a robust or insecure system depending on future demand. A scan across the planning horizon reveals a growing likelihood of running a yield deficit, with insecure projections increasing from 5 % of the total projections in 2025 to more than 75 % in 2055.

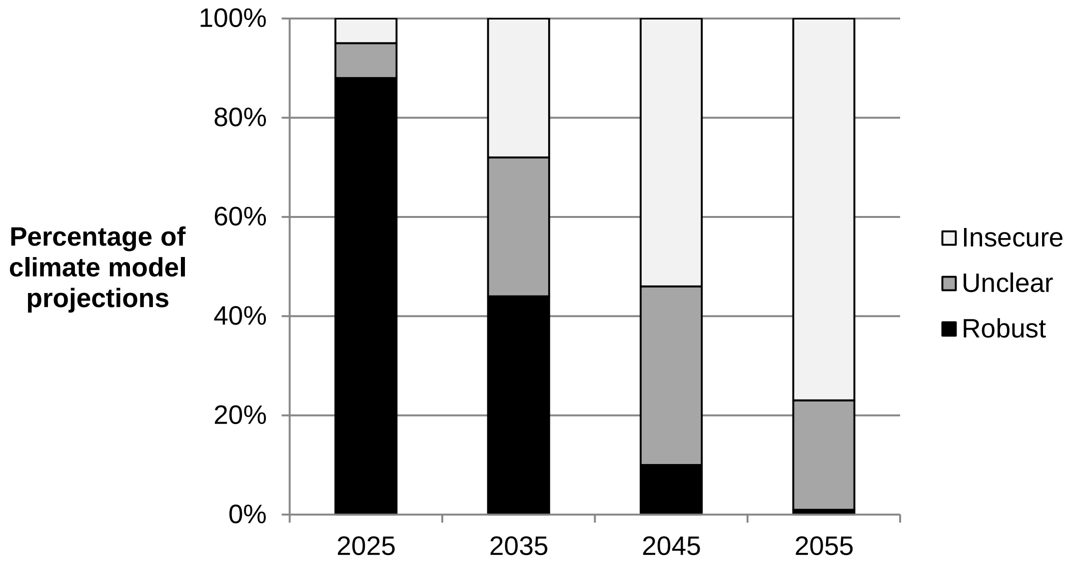


Figure 5-8 Proportion of climate model projections denoted the system as ‘robust’, ‘unclear’ or ‘insecure’ depending on their position relative to the decision thresholds.

The plots in Figure 5-7 indicate the existence of enormous uncertainty. This is caused primarily by poor agreement between the climate models on how a warming climate would affect mean annual precipitation. Figure 5-7d shows that by 2055 the MAP projection for the highest-impact emissions scenario ranges from approximately +2 % to -25 % from the 1990 baseline MAP. Figure 5-9 draws out this point more clearly, highlighting a distinction between the MAT projections, which embody relative consistency of change between the least and most severe emissions scenario, and MAP projections, for which agreement on sensitivity to emissions scenario is poor in comparison.

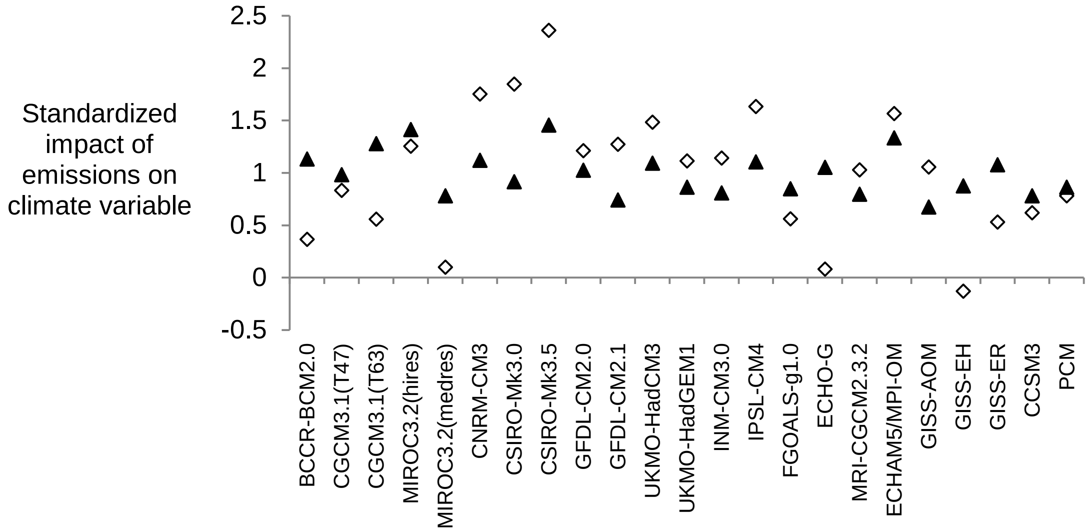


Figure 5-9 Difference between most (A1F1) and least (A1B) severe climate scenarios for 23 climate models, standardised by dividing by the mean difference across all climate models. MAT is represented by the dark triangles; MAP by the diamonds. The data are based on the 2055 climate projections from 23 GCMs (downscaled to Yarra ranges, set to change from 1990 baseline).

5.3.2 Flexibility of output

Figure 5-10a displays a filtered version of Figure 5-7b, plotting the 2035 projections from five climate models that were found to significantly outperform the others used in this analysis in a comprehensive world-wide assessment of GCMs and their ability to reproduce climate statistics under historic emission forcing (McMahon *et al.*, 2014). Specifically, the cited study used a Nash-Sutcliffe efficiency index and RMSE between modelled and observed annual temperature and precipitation statistics to rank model performance. The top five were established using this ranking in combination with a review of literature on the performance of the CMIP3 models.

Rather than focusing the projections in a single direction, these five GCMs delineate two distinct possibilities for reductions in MAP with warming climate. MICO3.2(medres) and ECHO-G predict a ~1 % reduction in MAP per degree of warming; UKMO-HadCM3, MRI-CGCM2.3.2 and ECHAM5/MPI-OM predict a ~7-9 % MAP reduction per degree of warming (Figure 5-10b). Importantly, either prediction taken alone would support fundamentally different planning decisions compared to the other. The 'dry' climates—clearly breaching the decision threshold even in 2035 (two thirds of projections beyond the expected demand threshold)—would create an impetus for additional supply sources within the next 20 to 30 years. But the climates that retain current MAP levels fail to seriously breach the decision threshold before the 2055 plot (referring back to Figure 5-7c, none of the 10 'wet' projections lie beyond the low demand threshold). These predictions might support a minimal-intervention strategy over the next 30 years. Given the investments required to provide water supplies to large cities and the uncertainty presented by these models, this analysis provides insights to water planners in understanding the level of risk exposure to climate change.

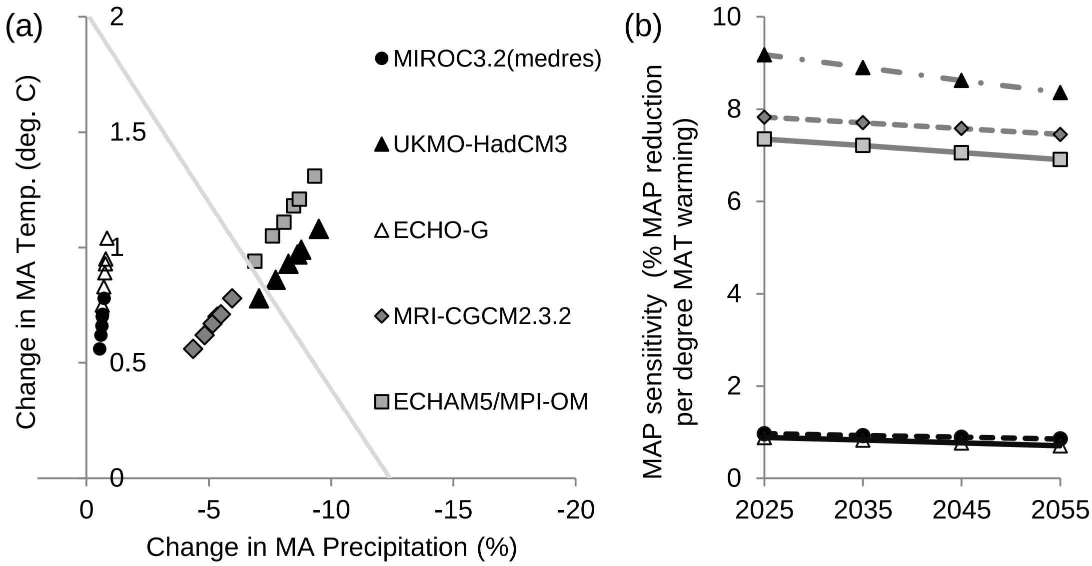


Figure 5-10 (a) 2035 climate impact analysis with data supplied from only the 'top five' climate models; (b) Two distinct sensitivities of MAP to warming temperatures, as projected by the 'top five' climate models.

Figure 5-11 further demonstrates the flexibility of the outputs of our analysis by displaying the data in a more classic 'supply-demand' format. The plot was achieved by back-calculating from the climate plots using the pre-derived climate response function, which returns yield for given MAT and MAP. The perforated black lines show the same expanding demand forecast uncertainty as mapped to the decision thresholds in Figure 5-7. The 2015 yield was back-extrapolated from the other climate years. The uncertainty shown here reflects that of the GCMs, which disagree even across the historical climates.

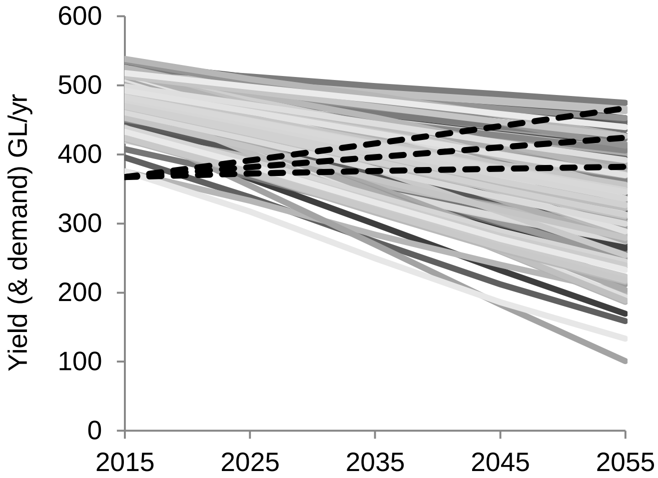


Figure 5-11 Results converted to the classic yield format: 138 yield projections (six emission scenarios × 23 GCMs) against demand (dark perforated lines).

5.4 Discussion

The results presented above demonstrate an instance of decision scaling successfully applied to a large and complex urban water resources system. The analysis provides a fresh set of insights around Melbourne's risk exposure to climate change. However, from the viewpoint of the urban water resources planning community, the important question is whether this version of the decision scaling approach could be applied more widely to assess climate risks on different municipal water supply systems. The following discussion therefore examines the assumptions and limitations of our approach to highlight potential issues for wider application. We also consider the extent to which the type of information produced by our analysis could inform difficult planning decisions around whether and how to invest in system improvements to expand capacity under climate and demand uncertainty.

5.4.1 Potential for wider application in urban water resources planning problems

The impact of water resources system complexity on the quality of the analysis presented above is brought to light in the strength of the relation between mean annual flow and yield. The analysis achieved a linear relationship with $R^2 = 0.74$. The remaining noise in this relation has been partially explained by the yield-constraining criteria; specifically, a given mean annual flow statistic could drive varying yield depending on the nature of the critical drought(s) in the sequence (Figure 5-4). We can therefore deduce that the applicability of the approach described in this paper would depend on the specific definition of yield in a given system. For instance, the ‘no-failure' yield is based solely on the critical drought and so the analyst may struggle to adequately relate it to long-term flow and climate statistics. On the other hand, it has already been well-demonstrated that flow statistics can predict a reliability-based yield in single reservoir systems using the Gould-Dincer equations (McMahon *et al.*, 2007). The work presented here suggests that these flow statistics can adequately predict a reliability-based yield in more complex multi-reservoir systems.

Rather than attempting to relate system performance to the net of the system inflows, we used the net inflows entering only four reservoirs in the system. Clearly, the other seven inflows entering the system would impact performance to some degree and a simple stepwise regression procedure could have identified the most appropriate set of inflow sequences to relate to yield across the 1000 replicates. Helpfully in this case, the four major reservoirs comprise more than two-thirds total system storage capacity and the other seven inflow sequences are well-correlated with our MAF-4 index, with correlation coefficients lying between 0.61 and 0.95. Other large systems may not exhibit this level of hydrological dependence and, in such cases, the analyst may struggle to successfully replicate our method. Nonetheless, our use of a flow index (MAF-4) facilitated the remaining steps of the analysis because the net of the mean annual flows entering the four-catchment area could be easily related to the available climate data and then to GCM outputs concentrated on the same region. This type of trade-off—sacrificing quality at one step of the procedure to simplify it in the next—could be the hallmark of bottom-up impact assessment of large, complex municipal supply systems.

Another potential limitation in this area relates to the definition of the regression line itself, which was drawn linearly in this study. We can deduce that the regression line would become non-linear and curve upwards at higher MAF values as yield becomes constrained by the capabilities of assets in the system. We can also deduce that a yield of zero (i.e., a system unable to sustain zero customer demand) would still require inflows to meet the environmental flow requirements in the system. The linear relation seems to capture this behaviour by suggesting that the MAF-4 should exceed ~160GL in order to generate a yield greater than zero. Yet our analysis failed to fully test the performance of the system at these lower MAF levels primarily because those statistics of the synthetic replicates do not deviate far enough from the historic statistics. We see this as a technical limitation of the synthetic data generation which has affected the outputs of this particular study, but note that the field of stochastic hydrology has already produced models that incorporate multi-decadal and nonstationary features that would overcome this issue (Sveinsson *et al.*, 2003; Koutsoyiannis, 2011; Salas *et al.*, 2012). Similarly, the analyst might simply alter the stochastic inflow model with a reduced mean to determine how the system responds to that particular climate change.

Similarly to Brown *et al.* (2012), we simplified our analysis by accounting for—and then setting to historic values—the standard deviation, skew and lag-1 autocorrelation of the indicator catchment annual flow series. This approach simplifies the remainder of the analysis by concentrating the climate impact assessment solely on the climatic changes that would affect the mean annual flow. But it imposes the assumption that changes in the other statistics will not be brought about by climatic change. Where necessary, this weakness could be mitigated by checking for the influence of changes in these statistics to provide additional rigor once a planning decision has been reached. Alternatively, the analyst could accommodate additional statistics—certainly the standard deviation of the annual flows—within the core procedure, as has been demonstrated elsewhere (Moody and Brown, 2013; Steinschneider and Brown, 2013).

We used a web-based service to obtain regional projections of climate change based on scaling of an observed baseline period. This method, and other scaling methods—e.g., the delta-change method (Nilsen *et al*., 2011) or quantile-quantile mapping (Vidal and Wade, 2008)—are straightforward to apply and can provide readily-applicable projection data sets for use in impact studies. However, when using scaled climate data it is important to remember that many of the underlying assumptions may not hold true under future climate change. Other more complex techniques exist for drawing out higher spatial resolution projections from GCMs. Fowler *et al*. (2007) comprehensively reviewed various techniques for the practice commonly referred to as downscaling. We note, however, that even though more sophisticated methods may have a theoretical ability to provide richer information of plausible change, these are associated with their own limitations (Feser *et al.*, 2011). For the Yarra Ranges, it is likely that the GCMs fail to adequately represent the regional processes that govern rainfall. GCMs operate on a course scale spatial resolution (~70-250 km) and will fail to capture the complex flow around the higher topography in this region. For this reason we might expect some bias in the projections used in this study and would expect similar problems to arise if our approach were applied in other regions.

5.4.2 Scope for improved urban water resources planning

One of the principal advantages of decision scaling is that it identifies system vulnerabilities that may otherwise not be discovered in a top-down approach. Whilst the uncertainties examined in a top-down approach are constrained by the climate projections, the more extensive vulnerability analysis characteristic of a bottom-up assessment may identify plausible risks that planners may wish to mitigate irrespective of whether projections breach the threshold of interest. At the very least, they would avoid developing an analysis that assumes the worst-case climate model projection matches the worst-case possible climate future.

Additionally, the climate response function enables the analyst to effortlessly update the assessment with new climate information, avoiding the need to translate climate projections into streamflow impacts. One potential advantage made available by this convenience—which is demonstrated in this paper—is the ease with which different sources of climate information can be separated and dissected (Figure 5-10). Stainforth *et al.* (2005, 2007) suggest that analysts could reasonably include or exclude different climate models based on their applicability to the specific problem such that the resulting understanding of risk might be more credible than that arising from use of multi-model ensembles. We used an assessment of a 'top five' models—albeit based on a globally-reaching rather than Melbourne-focused analysis—to predict two possible alternative pathways of change, each recommending a very different set of planning decisions. Given this choice, the planner might look to first avoid sunken investment and monitor and evaluate the changing climate to assess whether to switch strategy if and when the severe pathway seems more likely. A prior modelling study could determine whether existing operational flexibility (e.g., earlier triggering of supply enhancement or demand reduction options) could act as a cost-effective insurance policy to mitigate the risk of the severe pathway and provide a buffer during the time it would take to build new system capacity or further reduce demand should that future materialise.

Alternatively, the planner could design the system to satisfy a pre-determined percentile of climate scenarios described by the analysis. This approach would mirror the common practice for using 'top-down' derived uncertain impacts on yield (e.g., the 'headroom' approach prescribed by the England and Wales water resources planning guidelines). However, the question would still remain as to what is an appropriate level of risk coverage across the projections. We wish to emphasize that more detailed information describing possible future hazards and vulnerabilities does not necessarily enable or encourage more 'effective', 'robust', or 'successful' planning decisions, as suggested in recent literature (Dessai *et al.*, 2009; Wilby, 2010; Weaver *et al.*, 2013). This is not to ignore or censure the possibility for using that information in the ways described above or, for example, to inform public debate about the water security risks facing a particular region. But one should also consider that some planners may wish to avoid explicating the problem in a way that nurtures what they see as unproductive discourse on the difficult trade-offs between non-monetised drought-related risks and the substantial economic and societal costs of infrastructure development. The point is that an improved understanding of climate risk may only enable improved planning if the accepted decision frameworks can accommodate it.

5.5 Conclusions

We have developed and applied a yield-based decision scaling approach to assess climate impacts on the Melbourne water resources system. We believe this is the first such assessment on a large and complex municipal water supply system. The primary contributions of the paper are located in the evidence and discussion around the scalability of a decision scaling approach to large and complex municipal supply systems, and the critical analysis of the usefulness of our specific approach for assessing climate change impacts and then informing urban water resources planning decisions.

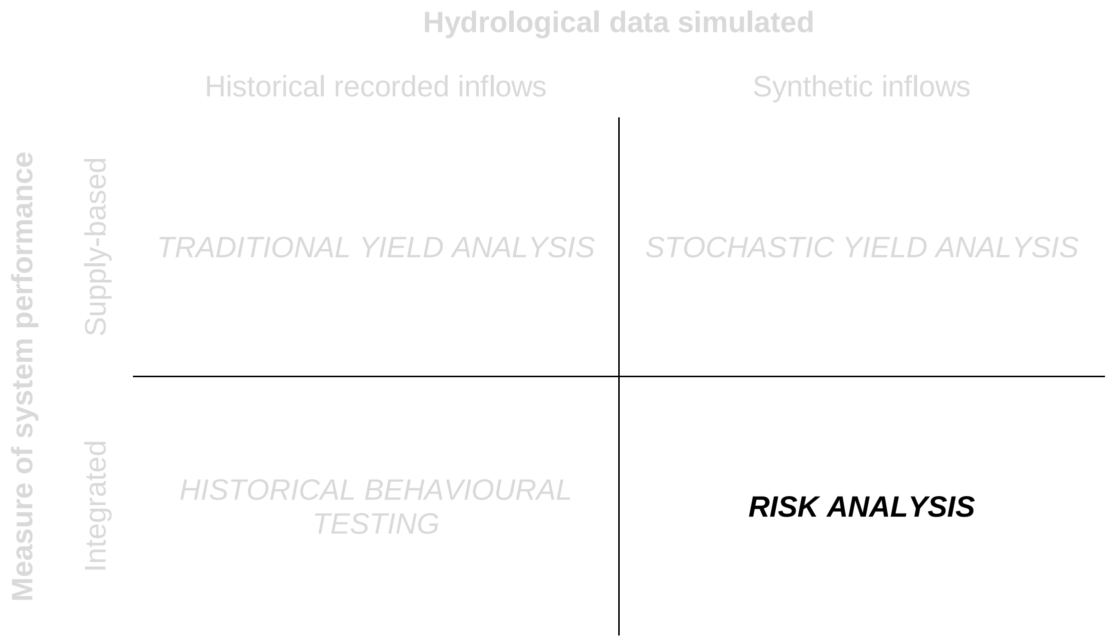
The strength of the relationships developed in our analysis—most importantly, that between mean annual flow and system performance—reflects the validity of the approach in the case presented. Our discussion considered the technical limitations and described a number of conditions under which the method might fail to provide adequate results. In particular, the quality of the analysis would depend on the criteria used to constrain yield, as well as the spatial extent of the system and the potential for indexing performance to a particular set of catchments within it. Nonetheless, we have provided strong evidence to support the view that a decision scaling framework can accommodate large and complex systems.

We measured performance using yield, which, contrary to performance metrics that integrate supply and demand (e.g., reliability), allows for the simple introduction of demand forecast uncertainty in a decision scaling procedure. Moreover, the use of yield allowed us to incorporate two separate service criteria into the assessment without compromising the clarity of the outputs, although we concede that the crucial MAF-yield relationship would suffer with the addition of more constraining criteria in the yield definition. We suggest therefore that the use of yield as a performance measure—which remains industry-standard—adds a new dimension to the small but fast-growing literature on decision scaling approaches. This development may be of particular interest to the urban water resources planner.

The picture of climate uncertainty generated by our approach could be incorporated into the modes of planning based on capacity expansion to meet a service threshold by (e.g.) picking a desirable proportion of climate projections against which the system would be designed to cope. The ease of updating, separating and dissecting climate projections would appear to open up opportunities for including or excluding different sources of climate information depending on credibility. For example, one might be interested to follow models that represent climate pathways that are of particular concern for a water security point of view as well as models that simulate the most likely change for the region. The selection process is non-trivial due to the sheer amount of data involved, but could be facilitated by tools such as the framework proposed by Whetton *et al.* (2012). Perhaps new research in this area should build on methods for assessing climate model credibility under different circumstances and develop ways for tracking the progression of climate through time to inform adaptive strategy underpinned by a decision scaling framework.

CHAPTER 6 STOCHASTIC MODELLING TO INFORM CAPACITY EXPANSION PLANNING

Article title	*Standardizing traditional water availability assessments by correcting the reserve storage bias*
Co-authors	Neil Upton (Atkins Global), Mark Smith (United Utilities PLC), Richard Blackwell (United Utilities PLC), Paul Jeffrey (Cranfield University)
Co-author contributions	Data provision, style guidance and corrections, general discussion.
Publication status	Ready to submit (Journal to be confirmed)
Additional notes	The ambition of this study differs significantly from the other three presented above (see section 1.4 for reasoning). It does not trial a planning framework that seeks to inform decisions under uncertainty. However, the study employs stochastic modelling on one of the largest and most complex WRZs in the UK. It thereby contributes to other aspects of the thesis relating to practicalities of stochastic modelling assessments.



6.1 Introduction

The concept of "yield" has underpinned water resources system design for more than a century. In capacity expansion planning of integrated, conjunctive use, municipal-supply water resources systems, a water authority typically schedules infrastructural investments (or demand management measures) to ensure that the system yield exceeds target demand over a given planning horizon (Loucks, 2005). Yield, in this sense, equals the maximum annual average volume of water that can be supplied by the system, subject to specified operational constraints (reservoir control rules, abstraction limits, asset capabilities, compensatory release requirements, etc.) and a seasonal demand pattern, without violating pre-defined failure criteria.

Definitions of yield vary depending on the criteria used to define failure. An analyst might specify failure by frequency, magnitude or duration of supply shortfall, and then determine the yield assuming either the historical recorded flows or stochastically-derived flow sequences. Failure might be defined using multiple criteria, including a desired "level of service" based on a maximum frequency reservoir trigger crossings that correspond to restrictions on customer water use (e.g., Erlanger and Neal, 2005; Environment Agency, 2012; Thames Water, 2013; Turner *et al.*, 2014). A much simpler metric—often termed "safe yield" or "firm yield"—derives from early reservoir storage theory: the storage is sized to meet the target release without total depletion of the active reservoir storage under a repeat of the drought of record. This approach informed much of the reservoir design in the United States by means of the Rippl mass curve method (Rippl, 1883; Archfield and Vogel, 2005) and has remained popular in practice (Ratnayaka *et al.*, 2009; Water Research Foundation, 2013), although Rippl's graphical technique has been superseded by computational behaviour analysis (simulation), which better accommodates the complexities of conjunctive-use water resources systems (Mays, 2010).

Design by "safe yield" suffers some basic, well-known flaws. First, the underlying analysis fails to adequately expose the implications of alternative designs on the frequency, duration and severity of failures, as noted by Klemeš *et al.* (1981):

> "The weakness of the [non-failure] approach becomes apparent if we realize that it does not distinguish between a case where a given storage would result in a failure in water supply lasting, say, a year with supply falling to 10% of the target and a case where a failure would last a couple of days during which the supply would have been reduced to 95% of the target."

In this paper we sidestep this particular issue—which has been dealt with comprehensively elsewhere (e.g., Hashimoto *et al.*, 1982a)—and instead focus on a second set of problems associated with design by yield, namely those relating to the use of short inflow records and a consequent approach for dealing with hydrological uncertainty: the “reserve storage” margin.

The earliest use of a “reserve storage” margin in water resources system design probably coincided with the recognition that the brevity of available inflow records meant that Rippl-informed designs would be vulnerable to relatively minor droughts. The engineers of the early 20th century recognized this problem and initially attempted to remedy it by synthesizing hydrological time series, first by (literally) shuffling the recorded flows (which were marked on a deck of playing cards) and later by developing sophisticated models that could generate statistically coherent replicate sequences (Hazen, 1914; Sudler, 1927; Barnes, 1954; Maass *et al.*, 1962). The original idea was that streamflow time series could be characterized as a stationary stochastic process such that a large enough sample would contain a near fail-safe design drought (Klemeš, 1987). But the premise for this approach unravelled with the discovery of long-term hydrological persistence (Hurst, 1951); there exists no well-defined envelope of hydrological variability or quantifiable upper limit on the severity of drought that a given catchment or set of catchments may experience. Well informed hydrologists and engineers would recommend designing water resources systems with conservative planning margins to provide additional robustness (e.g., Beard, 1965). The reserve storage margin is thus implemented in contemporary safe yield analysis (based on the drought of record) such that modelled “failure” is incurred not when the active storage (i.e., operationally-available storage) is depleted, but when X % of the active storage is depleted, where X is equal to the total active storage minus the proportion allocated to the safety margin. This drives robust design by lowering the computed yield or, what is the same thing, increasing the required system capacity necessary to ensure that yield meets target demand.

Application of this form of planning margin (also termed buffer storage, contingency storage, reserve capacity, conservation storage, reserve storage, emergency provision, etc.) in contemporary yield-based water resources system design is evident from planning documents of the United Kingdom (e.g., Department of Environment, 1996; UK Water Industry Research, 2012b; United Utilities PLC, 2013), Australia (e.g., Erlanger and Neal, 2005; State of Queensland Department of Environmental and Resource Management, 2010), the United States (e.g., New Jersey Department of Environmental Protection, 2011; Rush *et al.*, 2011; City of San Diego Public Utilities Department, 2013; Tennessee Department of Environment and Conservation, 2013; Catawba-Wateree Water Management Group, 2014) and elsewhere (e.g., Umgeni Water, 2013). Though highly subjective and potentially over-conservative, this form of ‘humility’ against hydrological uncertainty has served the economically-advanced countries well over the course of a century (Lins and Cohn, 2011; Stakhiv, 2011). It has also been vindicated by newly-discovered historical droughts—exposed through paleo-studies (e.g., Frick *et al.*, 1990)—and recent record-breaking droughts—such as that experienced in southeast Australia (van Dijk, 2013).

Nonetheless, it is worth asking whether this traditional planning approach meets the needs of a modern water supply industry, which may wish to *benchmark* and *compare* yield assessments for separate, adjacent, independent supply systems. Such unbiased comparison would be necessary to understand the distribution of water availability within a region or country, determine where to focus on planning and management activities, allocate capital investments effectively to improve service quality, and provide consistent information to the general public and wider stakeholders. This study exposes the potential for a bias in cross-system comparison that is caused by the use of the reserve storage margin. We explain why and how this bias arises and present a simple method for quantifying and correcting it. We demonstrate the method using two contrasting water resources systems located in England and then discuss possible implications for water resources investment planning.

6.2 Causes of the reserve storage bias

The simplest way to size reserve storage uses a set percentage—for example, 20% active storage assigned to reserve storage in all systems under investigation. This would clearly cause bias in situations where supply systems differ in their demands and in their abilities to draft water from alternative resources (e.g., boreholes, desalination plants, inter-basin transfers, etc.) during drought. The reserve storage ought to account for the various factors that influence inflow and demand; it makes no sense to provide 20% reserve storage in two separate systems if the resulting volumes would be depleted over significantly different time periods during extreme drought. The level of buffering against supply failure would be significantly greater in one system relative to the other and thus the resulting yield assessments (and system designs) would become skewed toward increased reliability in the system with the greater buffering capacity. One way of dealing with this issue to calculate for each system a reserve storage volume based on the demand under drought conditions over a set number of days. This approach is applied in England and Wales (UK Water Industry Research, 2012b) and we term the resulting reserve storage margin the 'demand-based reserve storage' through the remainder of this paper. The demand-based reserve storage equates to the customer demand plus any compensatory releases (environmental flows) minus the water available from local supporting sources.

The problem with this approach is that it fails to reflect within the reserve storage volume the potential for new inflows entering the system during drought (which will vary depending on drought conditions), particularly where those inflows can be harvested and then distributed from remote catchments. In other words, the reserve storage fails to incorporate the benefits of system diversity and integration—also termed ‘synergistic gains’ (Hirsch, 1977; Montaseri and Adeloye, 2002). We can demonstrate this omission with a simple example. Using a single 50-year streamflow series for a catchment feeding a reservoir in northwest England, we constructed a basic single-reservoir system (SRS) model with an assumed demand of 50 ML/d and reserve storage equal to 30 days’ demand (1500 ML). We then constructed a multi-reservoir system (MRS) using the same inflow record and two other records from catchments also located in northwest England, feeding three separate demands of 50 ML/d (Figure 6-1). The reservoirs in this system also included reserve storage equal to 30 days’ demand (total reserve = 3 × 1500 ML). Storages were then sized to provide a system yield equal to the demand in both systems—such that, at first glance, both systems would appear to be designed to equal reliability. The question, however, is whether the demand-based reserve storage in the SRS provides the same margin of safety as the demand-based reserve storage in the MRS.

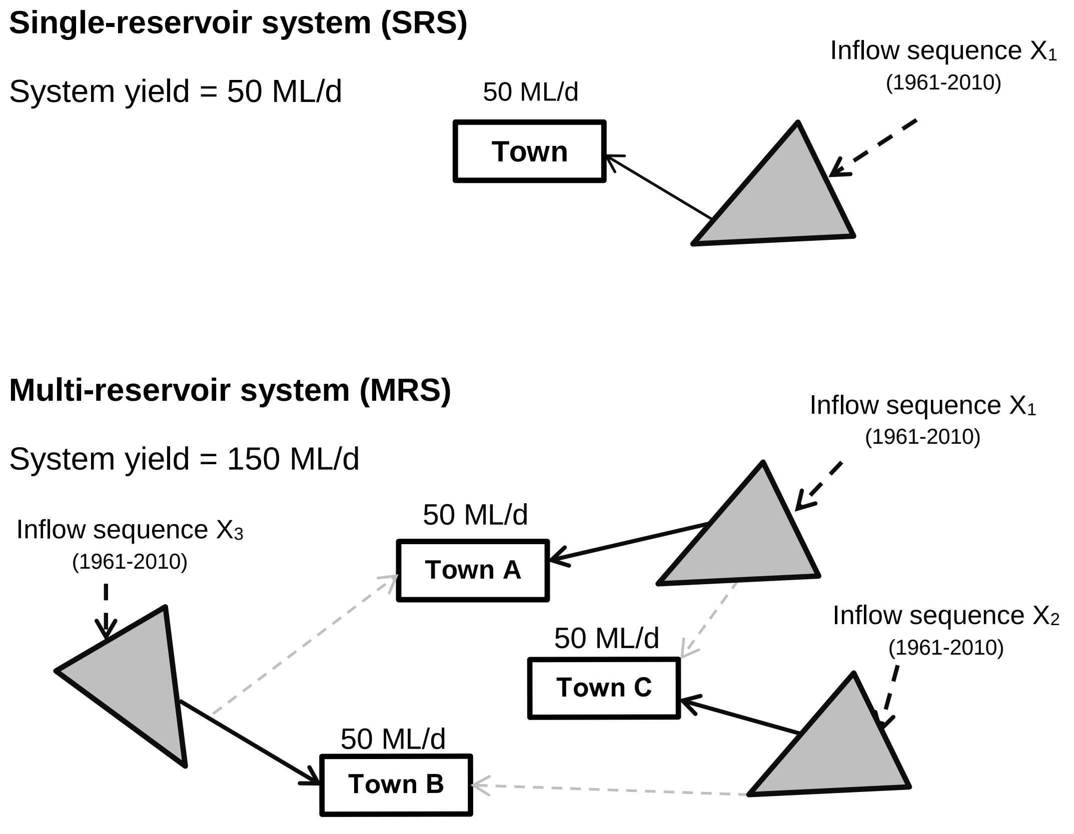


Figure 6-1 Simple single reservoir system and multi-reservoir system set-up.

Figure 6-2 shows the results of system simulations of four significant historical droughts. The potential for storage depletion in each system is indicated by the maximum of the 30-day moving average of total system storage depletion through the drought critical period as a proportion of the total demand-based reserve storage available in each system. The results indicate that the rates of storage depletion in the SRS would tend to exceed those experienced in the MRS when standardized by the demand-based reserve storage. This strongly suggests the reserve storage of the SRS would be more likely to suffer total storage depletion and coinciding supply failure. The yield metrics for these two systems, as currently configured, cannot be fairly cross-compared because the MRS enjoys a significantly higher margin of safety against failure through its reserve storage.

The other interesting point from Figure 6-2 is the variance in the standardized drawdown discrepancy across the different droughts. This suggests that the benefits of system diversity and integration depend on the characteristics of the drought under investigation—a point already noted by Hirsch (1977): “gains are a result of… the stochastic diversity of flows into each of the reservoirs.” In other words, the magnitude of gain in yield realized through system diversity and integration depends on the covariance of system inflows, which can vary from drought to drought. A rigorous approach to sizing the reserve storage margin must somehow accommodate storage behavioral patterns caused by a wide range of alternative plausible droughts. To condition the safety margin on a single drought may cause bias if the covariance of system inflows during that drought happens to depart significantly from the norm.

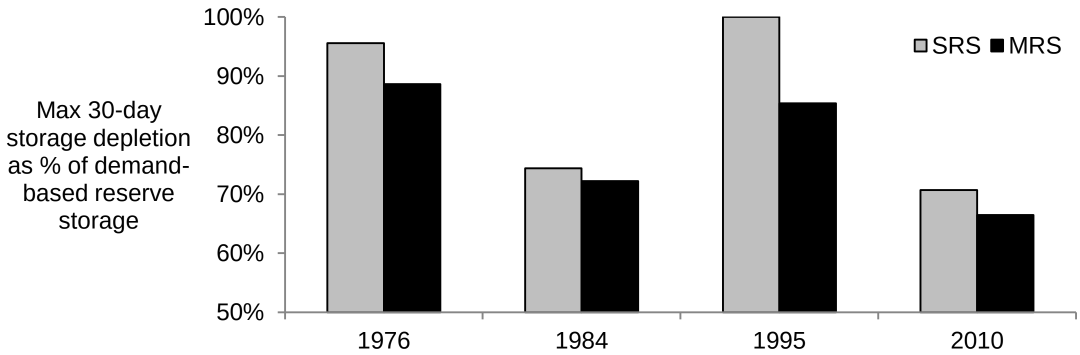


Figure 6-2 Maximum 30-day moving average storage depletion during four historic droughts that impact streamflows in the single reservoir system (SRS) and the multi-reservoir system (MRS).

6.3 A method for quantifying and correcting the reserve storage bias

6.3.1 Models and data

Models and data were obtained for two contrasting water resources systems located in England. These systems function as a test bed and communication tool for describing and reasoning our approach. The systems were stylized for the analysis and are termed System A and System B respectively. System A is a relatively small, isolated water resources system, supplying approximately 27 ML/d (annual average) water to a single town of approximately 100,000 people. The incumbent water service provider is licensed to abstract water from two local rivers, which constitute all of the supply in this system. Under specified low-flow conditions, river abstraction is partially constrained and water is drawn from a single reservoir, which is refilled during normal flow conditions. The reservoir provides an active storage of approximately 650 ML and would be depleted over a period of around three months under the critical historical drought conditions. System B is a relatively large and complex water resources system supplying approximately 1850 ML/d (annual average) water across a large metropolitan region containing numerous large towns, two major cities and a total population of nearly seven million people. The system exhibits a substantial level of integration and diversity. Water is primarily sourced from dozens of widely-dispersed impounding reservoirs (some more than 200km apart), and a significant portion of water is abstracted from local rivers and groundwater aquifers. A network of aqueducts and pipelines enables the transfer of water to augment diminishing storages and balance water shortage risk throughout the system. The critical drought of record would deplete the active storage over a period of approximately six to nine months depending on the level of demand.

Representative models of these systems were constructed using Aquator – a node-link mass balance simulator used widely in the UK (Oxford Scientific Software, 2008). The models were used to capture system specifications for bulk supply assets (storages, boreholes, river abstractions, inter-basin transfers, etc.) and their linkage to demand centres, to which seasonal demand factors were applied. Aquator enables users to implement various system operating rules through standard component features and a customizable code-based interface. For the systems described in this paper, rules were implemented to apportion reservoir compensatory releases based on storage levels, control abstractions based on license conditions, and initiate alternative supply sources and impose demand restrictions under low-storage conditions. More generally, and unless otherwise instructed, the model resolves the movement of water within the supply system using an optimizer that uses costs assigned to the various model components to minimize operational cost when sources are plentiful, switching to prioritize resource states when user-defined control curves and license use rates are breached. This particular software package runs on a daily time step and therefore requires daily inflow data. Relevant daily historical inflow records of 50 years' length were obtained (three inflow records for System A and 20 inflow records for System B).

Before proceeding, five historic droughts were simulated in the models in order to compare reservoir behaviour and indicate the potential for reserve storage bias. The simulations assumed current measured demands for water and results were extracted in the form of (active) total system storage (TSS) time series. Then, reserve storage volumes were sized for the storages in each system using the demand-based approach described above. Figure 6-3 shows average 30-day TSS drawdown rates experienced during drought critical periods as a proportion of the total demand-based reserve storage provided in each system. These results provide two important insights. First, the drawdown rates in system B, as a proportion of the demand-based reserve storage, appear to be less variable and on average lower (64% compared to 89%) than those observed in System A. This strongly indicates the potential for reserve storage bias caused by a failure to account for the benefits of harvesting and distributing new inflows entering the systems during drought. Second, the discrepancy between the two systems appears to be highly dependent on the particular drought under scrutiny. This suggests that a deterministic approach to sizing emergency storage (i.e., relying solely on the behaviour caused by a single drought) may fail to adequately capture the relative likelihoods of reserve storage depletion. In order to account for this variability, we go on to describe a probabilistic reserve storage sizing method based on a range of stochastically-derived droughts.

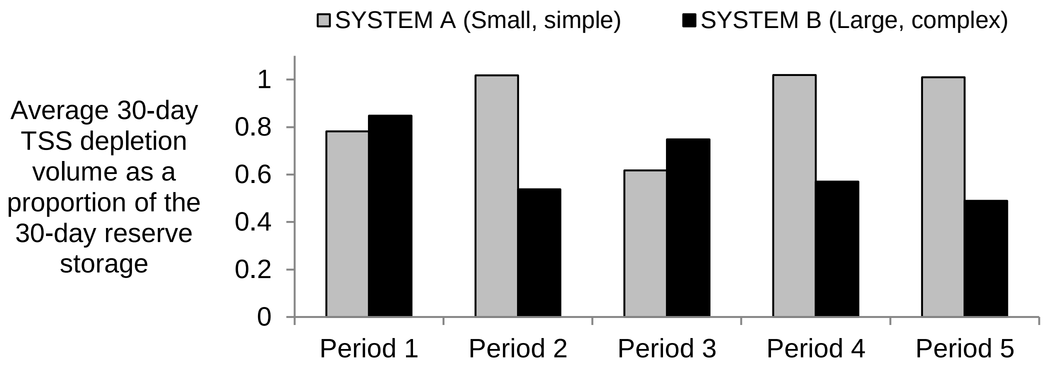


Figure 6-3 Comparison of average storage depletion volumes across five historic droughts that impacted both systems.

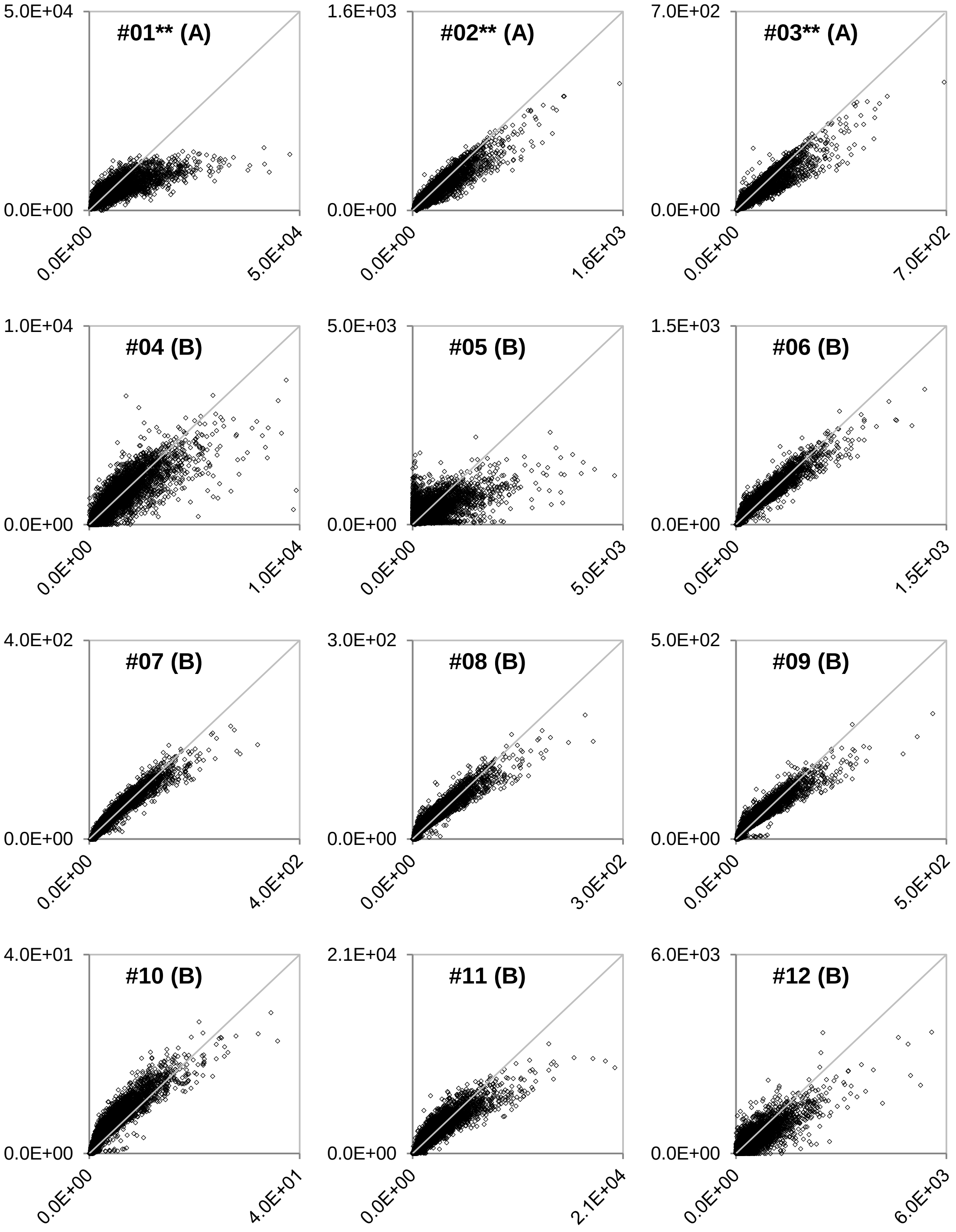
6.3.2 Data for a stochastic approach to sizing the reserve storage

A set of 23 precipitation records and corresponding evapotranspiration time series were prepared. These data corresponded spatially with the 23 catchment inflow sequences in the test systems. All 23 records were of daily temporal resolution and 36 years' length (1975-2010). Rainfall-runoff relationships were developed using an eight-parameter water balance model (Boughton, 2004). The model parameters were calibrated with a genetic algorithm in the 'Rainfall Runoff Library' (RRL) software package (Perraud *et al.*, 2003) to maximize a Nash Sutcliffe coefficient, using an exponent parameter to concentrate the calibration on low flows (method in Barma and Varley, 2012) for direct river abstractions where calibration to peak flows is inconsequential. Separate parts of the record were used to validate the models and, across the 23 models, mean Nash Sutcliffe scores of 0.768 (st. dev. = 0.132) and 0.759

(st. dev. = 0.148) were achieved for calibration and validation respectively. Graphical comparisons of simulated and observed flows are given in Figure 6-4. These fits were deemed adequate for the purposes of this study.

These rainfall-runoff models were used to produce long synthetic inflow sequences from daily synthetic precipitation and evapotranspiration data. In order to generate the synthetic rainfall data, a multi-site first-order two-state Markov model was applied (Srikanthan, 2005). (Data were synthesized from observed rainfall rather than flows to produce daily-resolution sequences.) The model generated 23 spatially-correlated synthetic rainfall sequences of 5,000 years' length corresponding to the 23 available records. Since this model deals solely with rainfall, a separate approach was needed to generate evapotranspiration sequences. A simple seasonal model (mean × monthly scaling factors) was calibrated for each site from the relevant evapotranspiration sequences, optimized for minimum root-mean-squared-error. These were then expanded out to 5,000 years' length to match the rainfall sequence lengths.

In order to accommodate the large datasets involved, the eight-parameter rainfall-runoff model was coded into R as part of an algorithm that would also read in the synthetic rainfall sequences, evapotranspiration data, and relevant calibrated rainfall-runoff parameters for each site. The code then generated the 5,000-year multi-site daily flow sequence set, pasted the data into a format that would be accepted by the resource system models in Aquator, and returned a range of statistics that would be compared with historic inflow data for quality control. Comparison of generated and historical sequences revealed excellent coherence for the first and second moments of the annual, monthly and daily statistics across the 23 sites (R^2>0.98 for all statistics). The measured covariance of the flow time series across the different sites was also well-preserved by the model. These data were then loaded into Aquator databases in preparation for behavioural testing of the water resources system models.


#01** (A)
5.0E+04
0.0E+00
0.0E+00
5.0E+04
#02** (A)
1.6E+03
0.0E+00
0.0E+00
1.6E+03
#03** (A)
7.0E+02
0.0E+00
0.0E+00
7.0E+02
#04 (B)
1.0E+04
0.0E+00
0.0E+00
1.0E+04
#05 (B)
5.0E+03
0.0E+00
0.0E+00
5.0E+03
#06 (B)
1.5E+03
0.0E+00
0.0E+00
1.5E+03
#07 (B)
4.0E+02
0.0E+00
0.0E+00
4.0E+02
#08 (B)
3.0E+02
0.0E+00
0.0E+00
3.0E+02
#09 (B)
5.0E+02
0.0E+00
0.0E+00
5.0E+02
#10 (B)
4.0E+01
0.0E+00
0.0E+00
4.0E+01
#11 (B)
2.1E+04
0.0E+00
0.0E+00
2.1E+04
#12 (B)
6.0E+03
0.0E+00
0.0E+00
6.0E+03

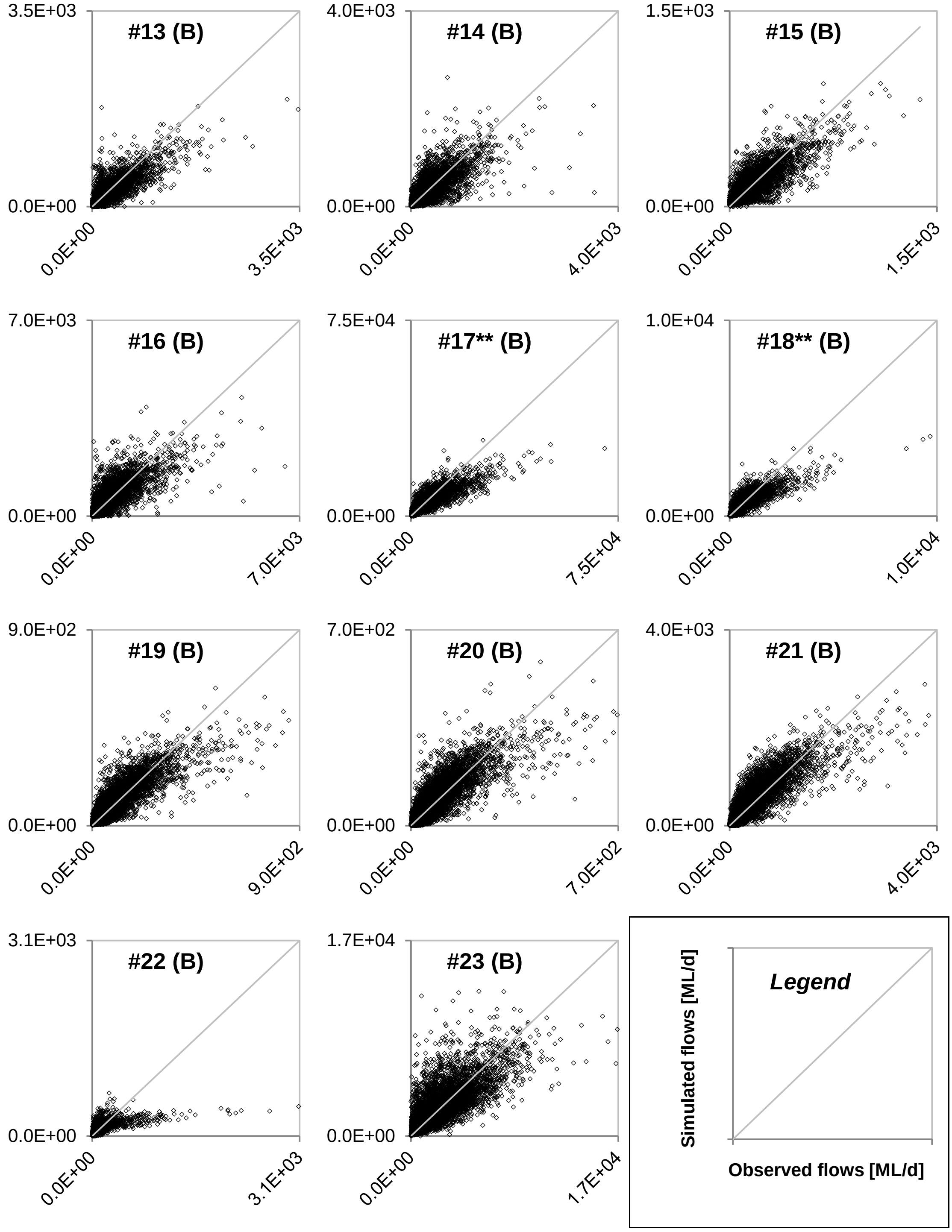


Figure 6-4 Graphical comparisons of observed and simulated daily flows for a 36-year record. Record names denoted with ** were calibrated with emphasis on the low flows.

6.3.3 A probabilistic assessment of system behaviour under critical drought conditions

Our approach for a consistent sizing of the reserve storage is based on the premise that the system models used to compute yield already contain the necessary specifications for understanding the system behavioural patterns that would be experienced during drought. These include: the initiation of emergency resources; demand restrictions; pumped water transfers; license volume breaches; staged compensatory releases to rivers; and—in the rainfall-runoff models—dry-weather catchment response. The influence of these factors will vary depending on the characteristics of the drought under investigation. The aim, therefore, is to simulate system behaviour under the range of hydrological conditions represented in the stochastic data and then size the reserve storage using a probabilistic understanding of the system behaviour, particularly total volume of available stored water that would be depleted over a given period of time. Drought conditions are the focus of attention because this is what is omitted when the reserve storage is removed from modelled storage in yield analysis. What is important is that the simulated system storage behaviour captures both the impacts of new water entering the system as well as the benefits of being able to distribute that water around the system under drought conditions.

A simple framework was developed to set a clear and consistent metric of TSS depletion during critical drought conditions for cross comparison between systems. Under this framework, the analyst begins by defining a specific drought occurrence interval – i.e., the L-year drought. The interval L should be large enough to ensure that the corresponding drought (i.e., the most severe event in the L-year sequence) causes severe storage depletion in the systems under investigation. Then, the synthetic inflow sequences are split into N sections of L years' length. Each section is simulated to capture the behaviour of the water resources systems assuming current demands and with all reserve storage constraints removed (i.e., the system is allowed to abstract from the lower reaches of active storage if required). The L-year critical drought is located in each section from the simulation results, which are expressed as TSS time series. The sample of N critical L-year droughts is then used to produce a distribution of storage depletion volumes that would be experienced during a specified X-day period leading to drought termination (Figure 6-5). By using the final period of the drought, the framework ensures that the analysis captures modelled behaviour most closely related to the behaviour one might expect whilst the system draws on the lower reaches of available stored water. The sample of X-day critical drawdown rates (and corresponding depletion volumes) forms the basis for a probabilistic understanding of system behaviour during severe drought conditions and can therefore be used to: (1) compare the relative margins of safety provided by existing demand-based reserve storage volumes applied across different systems; or (2) determine a more consistent reserve storage based on a percentile of the distribution of X-day depletion volumes, where both X and the chosen percentile would be held consistent across different systems.

This framework was applied to both System A and System B. The 100-year critical drought was selected (L=100) because the typical water resources system in England and Wales is designed against a 50-100 year historical inflow record; the 100-year droughts would therefore be expected to cause severe drawdown in most cases. The predefined 5,000-year multi-site synthetic sequence was therefore split into N = 50 sections, such that the analysis would produce for each system a sample of storage depletion behaviour under 50 alternative 100-year critical droughts. The separate sections of the synthetic sequence should not be considered 'replicates' of the historical data; clearly the use of a 5,000-year sequence split into sections longer than the available record prevents fair comparison of the synthetic drawdown behaviour with the historical equivalent. However, this poses no practical problem for the proposed methodology because the aim is simply to cross-compare critical storage behaviour across two adjacent systems experiencing the L-year droughts drawn from the same stochastic model. It is the cross-system comparison that reveals the bias rather than any comparison with recorded droughts.

For System A, the entire 5,000-year inflow sequence was simulated in the Aquator model and the resulting TSS time series was extracted. An algorithm was coded in R to first search each 100-year segment of the TSS time series for the critical drought, defined using the minimum TSS level, and then compute the average 30-day drawdown volume (ML) and rate (ML/d). The value X=30 days was selected in accordance with the current standard reserve storage in England and Wales.

An alternative approach was required to model System B because the complexity of the system imposes an impractical run time of approximately three weeks for a 5,000-year inflow sequence. In order to reduce the required computational time, the first 1000 years were simulated in the Aquator system and the corresponding ten 100-year droughts were located in each section. Characteristics of the corresponding net system inflows (i.e., the sum of the 20 inflows) were then used to define a simple model to predict the timing of the critical droughts. The critical drought location could be correctly identified in all cases by one of four moving average minimums of net system inflows (April-September; March-August; April-August; May-September). These moving averages were used to identify the possible critical drought locations along the remaining 40 sections of the synthetic inflows. An external procedure was written in VBA code to call and control the Aquator model. The code read in a potential critical drought location, opened the system model, simulated a three-year period containing the relevant critical drought, extracted the TSS time series, and then repeated under the next possible critical drought location. With this simple predictive model in place, the simulations were executed in less than two days.

The output of the probabilistic assessment framework comprises a distribution representing a range of possible TSS drawdown rates that would be experienced under severe drought conditions impacting both systems. We term this the 'probabilistic drawdown-based reserve storage'. To confirm the existence of an existing reserve storage bias between the two systems, one must simply relate these distributions to the demand-based reserve storage volume and then cross-compare.

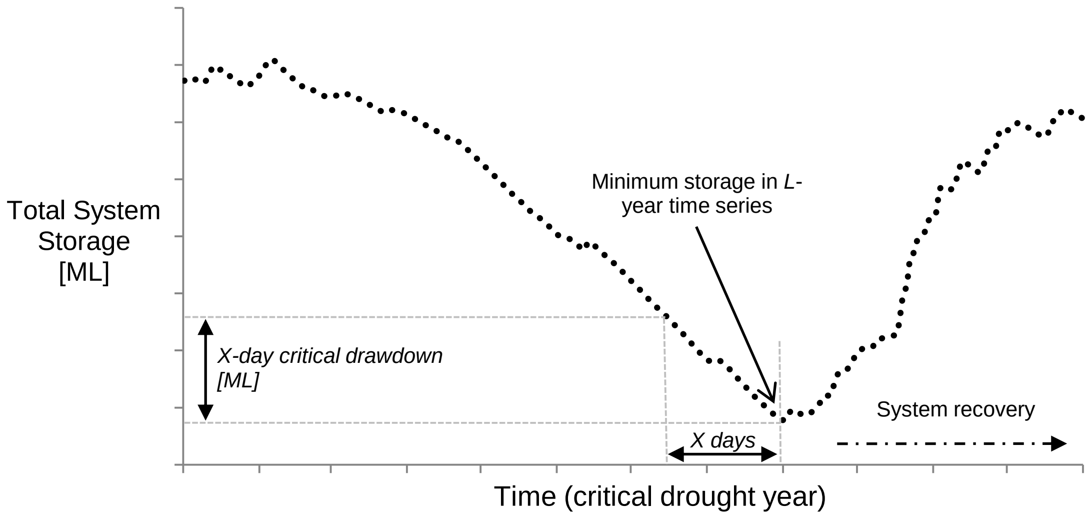


Figure 6-5 Positioning of the *X*-day drawdown rate in the *L*-year critical drought.

6.4 Results

Figure 6-6 and Figure 6-7 show the results of the stochastic storage depletion analyses for System A and System B respectively. The upper left quadrant on these figures shows annual TSS profiles for the 50 100-year critical droughts. A first observation is that none of the critical droughts exhaust the active storage, which is important because the method must avoid simulating conditions during which drawdown rates flatten due to lack of available water. System A exhibits behaviour typical of systems with low demand and low storage relative to catchment inflows. Such systems are characterized by short critical periods with substantial drawdown during short dry spells and fast system recovery initiated with relatively modest amounts of rainfall. As a result, severe drawdown and system recovery can occur at any point during the calendar year. System B, in contrast, exhibits characteristics of a system with large demand and high storage volumes relative to catchment inflows. Serious levels of storage depletion require prolonged spells of below-average rainfall, and relatively wet conditions are necessary to sustain a recovery. As a result, the critical droughts in System B tend to develop through the drier seasons with storages recovering in winter.

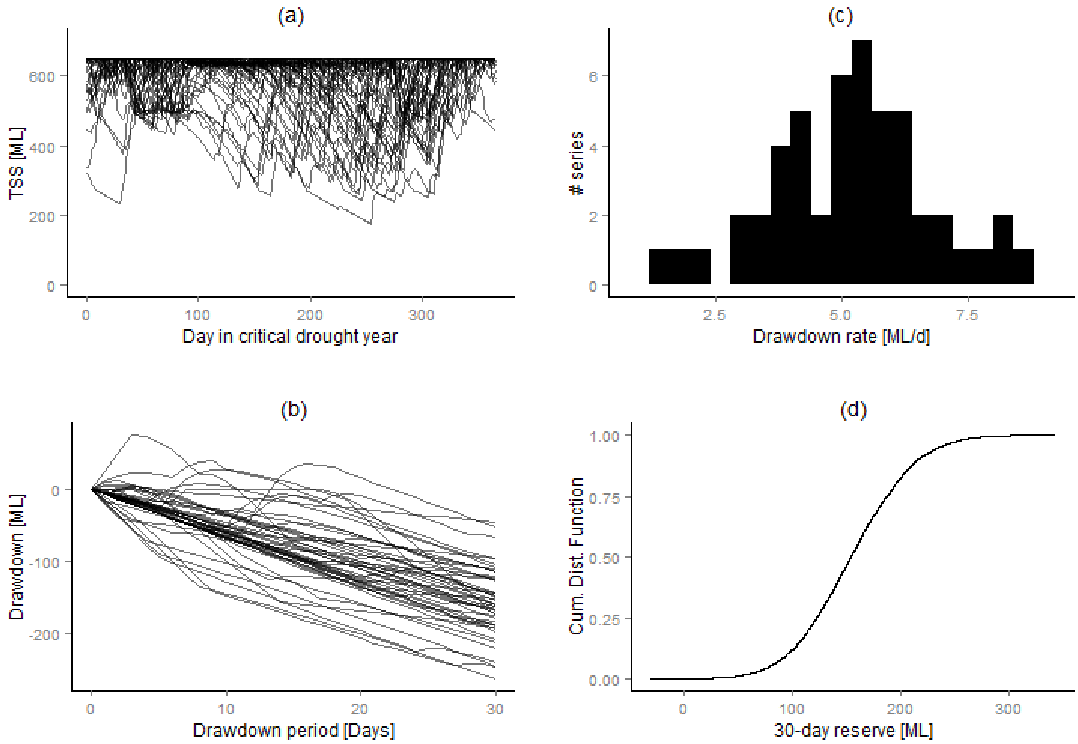


Figure 6-6 Stochastic critical drought drawdown analysis for System A

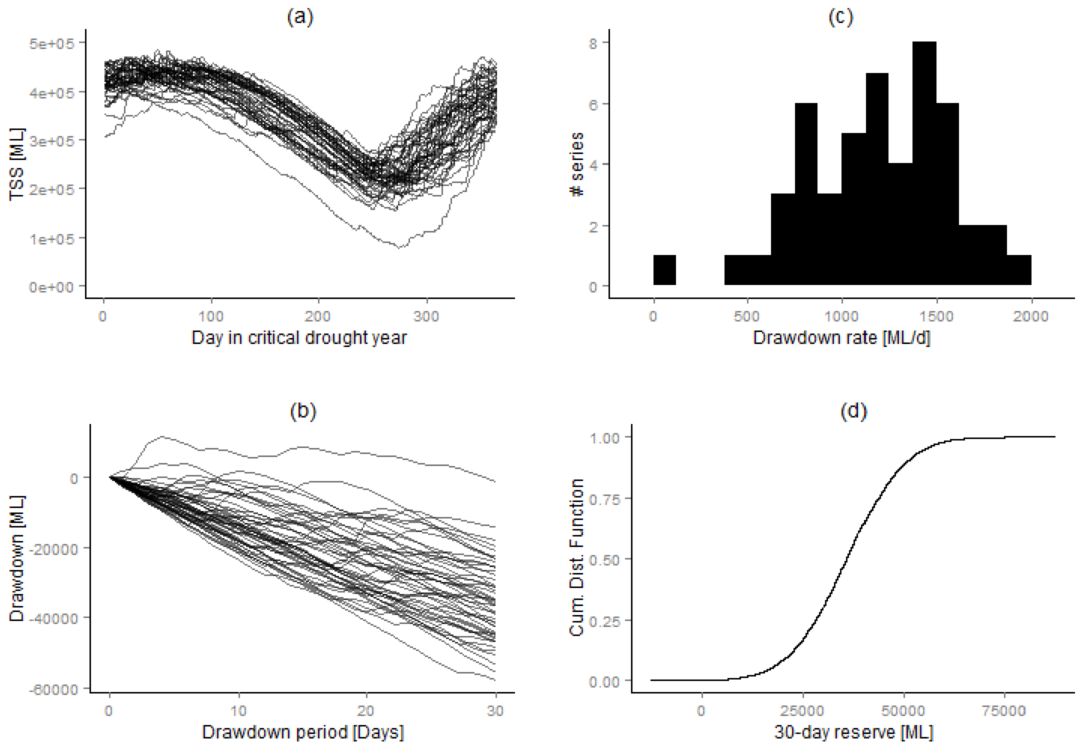


Figure 6-7 Stochastic critical drought drawdown analysis for System B

(a) Total system storage (TSS) profile for 50 synthetic critical droughts in their year of occurrence; (b) 30-day drawdown to termination of the drought (i.e. minimum storage point) from equalized 0 starting point; (c) histogram of average drawdown rates (ML/d) experienced during the 30-day drawdowns to drought termination; (d) implicit required storage to provide additional 30-day reserve.

Table 6.1 Synthetic (sample mean ± standard deviation) versus historic critical droughts

	System A		System B	
	Drawdown stats	*DBRS*	Drawdown stats	*DBRS*
Max. depletion [% TSS used]	47.2 (± 11.1)		55.6 (± 6.2)	
30-day crit. drawdown [ML/d]	5.2 (± 1.6)		1193 (± 385)	
30-day reserve store [% TSS]	24.0 (± 7.3)*	*31.0***	7.8 (± 2.5)*	*16.7***

based on critical drawdown rates multiplied by 30 days* * existing 30-day demand-based reserve store*

The lower left quadrant of Figure 6-6 and Figure 6-7 compares drawdowns during the final 30 days of the synthetic droughts. These depletion volumes form our comparable probabilistic metric of system behaviour during drought, as described above and depicted in Figure 6-5. The upper right quadrant presents these as drawdown rates (ML/d) in a histogram. Both systems exhibit a wide range of potential storage depletion rates during the 30-day period prior to drought termination. These results reflect the range of plausible changes in storage depletion that might occur during the final days of depletion of a system and thereby provide a probabilistic basis for examining and comparing existing demand-based reserve storage values. The lower right quadrant presents the required drawdown-based reserve storage volumes (rate × 30 days) as a normally-fitted cumulative distribution function. The data are actually slightly skewed, but the normal distribution will suffice in this instance to crudely demonstrate the scale of the discrepancy.

The results are summarized and compared with statistics derived from the relevant historical total system storage sequences in Table 6.1. It appears that the mean of the drawdown-based reserve storage values is significantly less than the corresponding demand-based value for both systems. Importantly, this discrepancy is far more pronounced in System B, where the mean drawdown-based reserve storage volume (equal to 7.8% total effective storage) is less than half the value of the demand-based reserve (16.7% total effective storage). We can further appreciate the differences between System A and System B in this regard by considering where the demand-based reserve storage value is positioned on the distribution of drawdown-based reserve storage values. In System A, the demand-based reserve storage maps to the 83rd percentile of the drawdown volumes displayed in Figure 6-6d. In System B, the demand-based reserve storage exceeds all of the modelled drawdown volumes displayed in Figure 6-7d (i.e., >99th percentile). This discrepancy between the two systems indicates a bias in the direction that would have been expected ahead of the experiments: by failing to account for the synergistic gains arising through inflow diversity and connectivity, the demand-based reserve storage provides a significantly larger margin of safety against supply failure in System B than in System A.

Figure 6-8a compares the two systems directly by standardizing each distribution of the drawdown-based reserve storage values as a function of the corresponding deterministic demand-based reserve storage. The drawdown rates in System B (per ML demand-based reserve store provided) are relatively low compared to those shown for System A, which reiterates the discrepancy noted above. We were able to quantify a bias correction factor by adjusting the assumed demand-based reserve storage to equalize these two distributions. We found that a 40 % reduction of the demand-based reserve storage in System B would achieve this correction—as shown in Figure 6-8b.

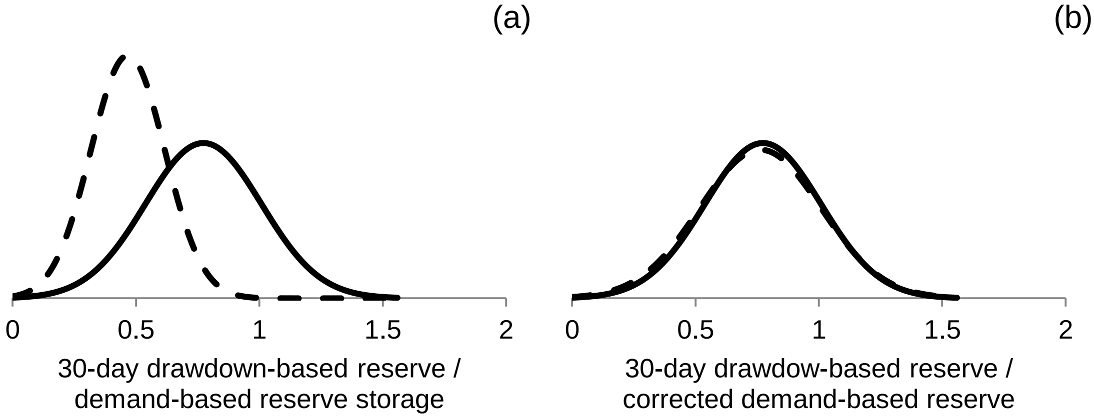


Figure 6-8 Probability distribution functions of 30-day drawdown volumes divided by the 30-day demand-based reserve storage volume for System A (solid black line) and System B (dashed line). (a) Unadjusted demand-based reserve storage. (b) Bias-corrected reserve storage, achieved by reducing the demand-based reserve storage in System B by ~40 %.

6.5 Discussion

6.5.1 Implications for water resources planning decisions

The most obvious implication of the results presented here is the potential impact of reserve storage corrections on system yield calculations and, therefore, understandings of water availability across different systems. We have calculated separately that the necessary 40 % decrease in reserve storage volume would drive a substantial increase in system yield in the order of 10-20 %. The reverse argument would apply if we wanted to increase the reserve storage of the System A to equalize with System B. In this case the increase in storage apportioned for reserve would substantially lower the system yield. These corrections are non-trivial and could drive significantly different investment plans where water authorities are responsible for multiple systems.

Water resources systems vary widely in terms of predominant source type, geographic extent, demand for water, local climate, diversity of inflows, and levels of connectivity. Despite this, we find only one water resources planning document that describes efforts to adjust for potential reserve storage bias, and none that use a structured approach to defining the necessary adjustments. Our study provides an indication of the level of reserve storage bias one might expect to arise across two specific contrasting systems. A closer examination of this feature across a wider range of systems would provide stronger evidence of the scale of the impacts of this bias in terms of yield comparability. However, a provisional conclusion would be that it is probably the case that for any two systems that perform equally well in terms of the water availability assessment ('supply-demand balance'), the more diversified system will likely be significantly better protected against supply failure through its reserve storage margin and therefore more reliable. This is not obvious from the outset because planners would expect the water availability assessments to fully account for the nuances of the different system designs.

6.5.2 Impetus for quantifying and correcting the reserve storage bias

Whilst the analysis presented in this paper clearly highlights the potential for problems when cross-comparing different supply systems, we ought to consider the reserve storage bias amongst other potential sources of bias to ask whether it would be worthwhile implementing a correction factor using a probabilistic drawdown-based reserve storage analysis. Furthermore, having identified the reserve storage as a potential problem, we should consider the consequences of simply removing it to improve the prospects for fairer water resources system comparison.

Other potentially important sources of bias in water availability assessments include: control curve positioning, assumptions on the limits of available reservoir water (constrained by pumping capacity or water quality conditions at lower reaches), the use of the critical drought from short historical records (the critical drought return period used to assess yield on one system could differ significantly from that used on another), and differences in uncertainty arising from flow gauging and naturalization. Planners may be interested in determining the relative contributions of each of these aspects to the overall bias in their water availability assessments. Given the scale of investments required to improve water resources infrastructure, we would assert that any practical effort to determine and then reduce these biases ought to be of real value to decision makers as they attempt to define a water resources plan that considers the balance of risk across their systems. So the presence of other source of bias should not deter efforts to quantify and correct for reserve storage bias.

The purpose of the reserve storage is to provide a margin of safety against the prospect of a drought more severe than the critical drought of record. One way of achieving the same goal – without implementing reserve storage – would be to base the yield assessment on an extreme critical drought derived from synthetic hydrology. This could be done by either selecting the most extreme drought from a very long sequence of, say, 10,000 years' length, or by using a distribution of L-year critical droughts and designing the system for given percentile from that distribution – e.g., the 95th percentile from a large sample of 200-year droughts. Such an approach need not be computationally demanding in the long term because the identified extreme 'critical drought' could simply substitute the historic sequence in the existing form of yield assessment (with reserve storage removed). It may be the case—as was found here—that a stationary stochastic model will fail to generate droughts of an extremity great enough to build a level of robustness commensurate with the in-situ reserve storage margin. Stochastic generators that incorporate nonstationary features (e.g., Sveinsson et al., 2003; Koutsoyiannis, 2011) or climate model ensembles (e.g., Prudhomme *et al.*, 2012) may prove useful in such cases. On the other hand, the planner may prefer to communicate the yield assessment in such a way that concretely and credibly demonstrates adequate performance through the critical drought of record. In such cases the reserve storage will remain an important planning heuristic.

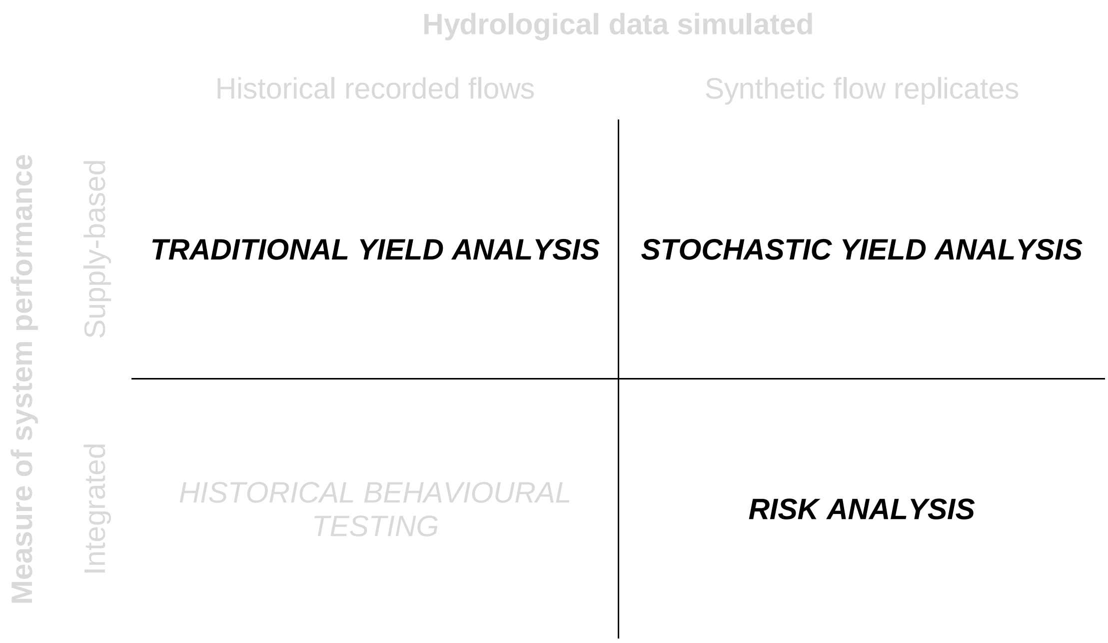
There are further practical challenges that planners might face if implementing the drawdown-based approach described in this paper. For example, the use of total system storage may not appropriately indicate the range of storage depletion volumes in more weakly-interlinked systems. In these systems some demand centres may be supported by isolated reservoirs that suffer rapid drawdown during drought, which may not be easily augmented from elsewhere in the system. In these circumstances the planner may struggle to adequately apportion the total reserve storage volume amongst the different reservoirs to provide equal safety margins throughout the system. Nonetheless, this paper demonstrates significant scope for improvement in the comparability of critical drought yield assessments that incorporate reserve storage. We believe our work could be advanced with additional research.

6.6 Conclusions

We have presented a simple probabilistic method of water resources system behaviour assessment that can be used to expose and then correct a significant bias in traditional water availability metrics based on yield analysis. This bias is caused primarily by differences in system inflow diversity and the resulting synergistic gains in yield, which can be captured in water resources system simulations but are generally overlooked when sizing reserve storage. As a result, planning authorities cannot fairly cross-compare yield-based water availability assessments, which will be important for water planning authorities that are responsible for multiple systems and seek a balanced picture of reliability in order to define their investment plans. Further research across a wider variety of system types would help build evidence as to the generality of these findings.

CHAPTER 7 INDUSTRY PERSPECTIVES

Article title	Industry views on water resources planning methods—prospects for change in England and Wales
Co-author(s)	Paul Jeffrey (Cranfield University)
Co-author contributions	Style guidance and corrections, general discussion, interview guidance.
Publication status	Accepted for publication in *Water and Environment Journal*



7.1 Introduction

Water resources planning practice in England and Wales has undergone significant change over the last 25 years. It has responded to the political, legislative and regulatory environment to become a formal, standardised process. It has amalgamated old and new concepts, and crystallised them in a consistent and clearly defined set of metrics and methods that form the building blocks of the Water Resources Management Plan (WRMP)—now a statutory reporting requirement for all water companies. Yet none of the methodological changes experienced since industry privatisation represent a fundamental shift away from classic design principles that have persisted for more than a century. Ever present features include: an assessment of water availability under a given set of conditions and assuming historical recorded inflows, which produces a deterministic metric of supply known internationally as "system yield" and today in England and Wales as "Deployable Output" (DO); the use of subjectively defined planning heuristics and engineering margins; and the communication of the state of the system using deterministic terminology, such as "surplus" and "deficit", which can also be quantified as the difference between yield and demand—the so-called "supply-demand balance."

Under these design principles, an impetus for investment is created if current or projected conditions open up a deficit between supply and demand. The aim is then to discover the least-cost combination of options to address the deficit over the planning horizon. This design paradigm is broadly known as "least cost capacity expansion" (Loucks *et al.*, 2005). Its weaknesses from a technical decision making perspective, which have long been recognised and are well documented, derive from its deterministic outlook amidst hydrological uncertainty and other uncertainties, which limits prospects for weighing the costs of infrastructure development against the benefits in terms of reduced risk. Capacity expansion planning contrasts with "risk-based" forms of planning that integrate supply and demand in the modelling assessments, explore hydrological variability and uncertainty using stochastically-derived flows and extensive Monte-Carlo analysis of the system, and produce likelihood estimates for a range of undesirable drought-related hazards.

A debate on whether to shift toward this form of planning has begun in England and Wales, where a number of recent academic papers have promoted new planning frameworks that purport to deal more effectively with uncertainty (e.g., Hall *et al.*, 2012a; Hall and Borgomeo, 2013; Korteling *et al.*, 2013; Matrosov *et al.*, 2013a,b). More generally, there is a view from academia and consulting firms that the climate change adaptation agenda demands a fresh approach for dealing with uncertainty in water resources planning practice (CH2MHill, 2013). This contrasts with the more conservative ambitions of the water companies (same study), which, in general, seek changes that would simplify and clarify existing process. So far this discussion has suffered from a lack of well-evidenced, documented knowledge to explain why the opinions of academics and practitioners diverge so starkly. This study aims to address this knowledge gap by capturing practitioner perspectives on the role of modelling and analysis in water resources planning. It seeks to understand what practitioners want to achieve through their modelling assessments and to expose the relevant viewpoints to help guide future research and inform policy decisions on how to shape the WRMP guidelines for 2019 and beyond.

7.2 Interview approach

The study targeted experienced practitioners with working knowledge of the water availability assessment methods prescribed by the WRMP guidelines (Environment Agency, 2012). Our focus on technical methodological aspects distinguishes the study from earlier research (e.g., Davies and Daykin, 2011) that has elicited practitioner views on the more general and administrative aspects of the WRMP process. The study is also distinct from the "Manual of Source Yields" UK Water Industry Research project and the aforementioned work examining the case for change in the planning guidelines (CH2MHill, 2013) because it seeks to explicate some of the underlying motives that shape practitioner opinions on planning methodology.

The research sought to answer three questions:

1. How do some of the accepted technical methodological weaknesses of DO assessments affect the ability of companies to plan effectively?

2. Why might practitioners seek to uphold the conventional modes of planning?

3. What institutional factors might hinder the prospects for a risk-based approach informed by stochastic modelling assessments?

We interviewed fifteen practitioners across seven water companies, the Environment Agency, Ofwat and two consulting firms. The sampling of participants was non-random; practitioners were selected for interview based on their role, experience and level of involvement with recent regulated planning activities. A small number of participants were recommended by other interviewees as appropriately experienced to contribute to the study. The sample incorporated a mix of lead planners (e.g., Head of Water Resources, Supply-Demand Manager), modellers (e.g., Modelling Manager, Water Resources Analyst/Consultant) and regulators (e.g., national-level co-ordinator, Regional Officer). All prospective interviewees were approached via email.

The interviews were conducted face-to-face, which allowed for the use of visual aids and sketches for explaining and discussing abstract ideas. The interviews were "in-depth," comprising pre-determined open-ended questions (i.e., "Why…?" and "How…?" questions rather than questions that can be answered with yes or no) and impromptu follow-up questions designed to extract as much information as possible on each subject (Foddy, 1993). A structured script was followed to prevent the discussion from drifting and to allow for responses to be categorised and compared. We used hypothetical scenarios—presented using water resources system schematics and performance graphs—to develop shared understanding and ensure subsequent questions were understood as intended. The script also included prompt information, including quotations from industry reports and reasoned arguments; the intention was not to lead the participant but rather to evoke counter-arguments and opinions on contentious issues, specifically those relating to weaknesses in DO assessment methodology (the approach was to first establish whether the interviewee acknowledged the existence of a particular weakness and then to elicit views on how the issue might affect a company's ability to plan effectively). A basic interview script was sent to participants in advance to give them the opportunity to understand the interview themes and raise any questions. All participants were advised that their answers would be reported in anonymous form, with the proviso that distinctions might be drawn between different groups of interviewed practitioners (i.e., regulators versus company planners). The typical interview lasted between 60 and 90 minutes.

Results are reported in the following sections using non-quantified terms—one, a few, some, most, almost all, all, etc.—to deter the reader from inferring proportional industry-wide representation from the relatively small sample of participants. All participants are considered to be 'practitioners.' Company practitioners are termed 'planners.' Practitioners from the Environment Agency and Ofwat are termed 'regulators.' Practitioners from consulting firms are termed 'consultants.'

7.3 Practitioner perspectives

7.3.1 Establishing weaknesses in DO assessment methodology

Three separate hypothetical scenarios were presented to elicit views on some possible weaknesses in DO assessment methodology. The scenarios were fabricated, but had been set up using real reservoir inflow data and modelled using Aquator (Oxford Scientific Software, 2008). The first scenario was designed to highlight a problem of capturing *vulnerability* using DO assessments. It featured two hypothetical single-reservoir Water Resource Zones (WRZs): one was in a state of DO deficit (DO = 95% demand approx.), but was supported by a desalination plant that would be able to supply a third of the demand if the reservoir failed; the other was in DO surplus, but was isolated such the zone would suffer total demand shortfall if the reservoir failed (i.e., the zone in deficit was less vulnerable than the zone in surplus). Participants were presented with supply-demand balances and supporting risk analyses to highlight the state of both systems. They were then asked in which of the two systems they would invest an arbitrary sum of money for alleviating risk, ignoring regulatory guidance (which would direct investment to address the DO deficit) and assuming only one system could be selected for improvement. Very few participants —and no planners— argued that the system in deficit merited the investment. Roughly half of the participants reported that the system in surplus merited the investment, in most cases commenting that the consequences of failure in that system would be unacceptable. Most of the other participants reported that they would be uncomfortable reaching a decision without further information, particularly customer preferences.

The second scenario focused on the issue of *subjectivity*. The scenario showed how a reservoir system could shift from a state of significant DO surplus to significant DO deficit depending on the size of the Emergency Storage, which is influential if DO is constrained by critical drought failure. Participants were asked for their interpretation of the reasons behind industry rules for sizing and reporting of the Emergency Storage. Most participants commented that the Emergency Storage margin came out of the Government's Agenda for Action (Department of Environment, 1996) that followed the 1995/96 drought, but none knew why 30 days' demand is the recommended volume. Some participants guessed that the assumption had been based on an assessment of the time required to either ensure supply into wetter autumn months or to implement contingency measures. Others believed the reasoning was more superficial: 30 days is a "nice number," equal to a calendar month. But most either gave no reason or stated explicitly that there is no logical reasoning behind the volume used. A few participants reported that Emergency Storage is too difficult an issue to tackle properly and, as such, it has never been challenged in a risk management context. Moreover, the majority of participants believed that the omission of Emergency Storage in guidelines and reporting requirements is an oversight—that the Emergency Storage "has slipped through the net", is "lost in history", and has been "buried in the analysis and forgotten about." One planner described a "tick-box mentality" that has led regulators to overlook important details, including Emergency Storage. A few participants commented that such arbitrary assumptions lead to a weak understanding and prevent clear communication of risk to customers.

A third scenario focused on *comparability* of DO assessments. The aim was to highlight how certain system characteristics (e.g., level of integration and source diversity) might prevent fair comparison of DO assessments across different zones. The scenario turned out to be superfluous because the idea that DO assessments cannot be fairly cross-compared across WRZs was already an accepted fact in nearly all of the participants' minds. Most reported that fair DO cross-comparison was an unrealistic goal and that there are numerous causes of inconsistency across different WRZs, including emergency storage, source types and control curve positioning.

7.3.2 Implications for effective planning

The basic scenarios and subsequent discussions established wide practitioner agreement that DO assessments (1) may fail to capture problems of resilience and vulnerability, (2) are often influenced by highly subjective assumptions, and (3) cannot generally be compared across different WRZs. Unsurprisingly then, nearly all of the practitioners reported that there was a case for a more detailed exploration of risks in the guideline methodology. Several participants suggested there was a need to extend the understanding of resilience and the consequences of system failure. Others argued that there was a need to better communicate the risks, costs and trade-offs in the plans. One participant reported that there was a need to recognise the value of investments that extend supply-demand surplus. Moreover, several participants added to the critique of existing DO assessment methodology by identifying some additional problems, including: a lack of any requirement to estimate likelihoods for a repeat of the drought of record or for a drought that would cause catastrophic failure; a failure to adequately expose the risk trade-offs between different interest groups; a lack of any requirement to understand and expose the "real consequences" of drought (e.g., would standpipes really be implemented at the lower storage triggers, or would a state of civil emergency generate sufficient political impetus to begin breaching environmental thresholds to uphold customer supplies?); and a failure to capture "robustness"—the capability of a system to uphold required performance standards under many alternative plausible futures.

Yet despite all of the issues identified above, most of the participants argued that the existing process is imperfect but adequate; only a small minority saw any need for fundamental changes in the planning process. The identified weaknesses (vulnerability, subjectivity, inconsistency, etc.) were not widely regarded as a major impediment to effective planning within the current framework. The general line of thinking here was that DO assessment, despite its fallibilities, does not necessarily constrict a company from accommodating other sources of information in its plan; the overall process allows flexibility for important elements to be incorporated even if DO assessment fails to expose them. This view was typified by one participant's claim that existing process is not "black and white"—it can accommodate political influence, the experience and knowledge of planners, or even additional analysis. One practitioner cited published planning appendices to demonstrate that certain companies had investigated resilience and vulnerability separately as part of their plans. On the issue of subjective margins, many participants emphasised that the aim of the prescribed planning process is not to attempt to define the "right" level of protection, but to define a baseline against which alternative options and scenarios can be tested. These practitioners argued that the primary function of the planning guideline is to provide transparent and consistent principles, minimum technical standards, an easily-understood basis for asking questions and a systematic, simple and repeatable method. Few saw any need for the baseline DO assessments to be comparable across the different water companies (although nearly all believed that a company should be able to cross-compare the water availability assessments across its own WRZs). The general feeling was that consistency of method and principles was much more important than comparability of the output metrics.

Some participants also discussed ways in which to improve current practice without shifting away from DO assessment and capacity expansion planning. For example, a few participants suggested that the Emergency Storage margin could be removed and the underlying uncertainty brought into Headroom to remove some of the subjectivity and improve transparency.

7.3.3 Effective communication

Participants were asked what possible problems might arise from communicating risk in a simplistic way through the supply-demand balance (the question was primed with a statement noting that customers may assume that "surplus" indicates zero likelihood of failure, or that "deficit" indicates extreme risk). A few practitioners reported that stakeholders can be easily misled by the supply-demand balance—it "hides complexity," "creates a blinkered view of what's going on" and creates a false impression that "right level of risk is known." One participant reported that politicians tend to get an oversimplified impression that a certain level of investment can eradicate risk, which fosters distrust in the industry if and when the fragility of a system is exposed by severe drought conditions. Another noted that the supply-demand balance fails to present the cost-benefit trade-off. One participant suggested that term "surplus" might create a misleading impression that water that can be traded away without significant impact on risk.

Most practitioners spontaneously attempted to justify the need to communicate the state of a WRZ using a supply-demand balance. Many argued that the supply-demand balance is the only way to get an informed response from customers. The dominant belief was that companies need a simple and "clear" way of alerting customers to any issues, and that the best way to achieve this is with a pass/fail test and the term "deficit." This perspective re-emerged at a number of different points during the interview. For example, when asked why classic design principles have persisted, a number of practitioners focused on the need for a clear pass/fail test, because "customers will tend to agree that a deficit needs to be addressed." One practitioner captured the general sentiment by reporting that the supply-demand balance is more important as a communication tool than as a means to reaching decisions on where and how to invest in system improvements. Similarly, when asked about the overall purpose of the planning process, a number of participants focused on the value of prescribed guidance in terms of the legitimacy it creates for helping build the case for investment. They reported that the main purpose of the guideline is to expose risks, to identify and explicate the need for new resources and, ultimately, to "help companies invest to protect supplies for customers."

7.3.4 Prospects for risk-based planning

It appears that an important advantage of DO assessment is its deterministic output, which helps planners communicate risk in simple terms. It should follow then that a method that produces complex metrics that are difficult to communicate would be deemed unattractive. Indeed, the interviews confirmed this supposition. Several participants reported that stakeholders would struggle to understand probabilistic performance metrics. Some feared that exposing difficult risk-cost trade-offs might open up non-productive debate on whether and how to invest. One participant used the term “paralysis by analysis” to describe this risk; another stated that investment based on probabilities would be a “hard sell.” One participant also warned that stochastic hydrology would cause confusion and emphasised the importance of using DO based on the drought of record, which provides clear assurance to customers that the system is designed against a real and tangible event.

Several participants questioned the validity, rigor and worth of risk-based planning frameworks from a decision making perspective. For example, a third of participants indicated distrust in the plausibility of synthetic droughts generated by stochastic models and suggested that sensitivity analysis of the drought of record would be a more rigorous and appropriate way to examine risk and uncertainty. Some practitioners also identified potential problems with defining the “right level of risk” and others commented that companies would struggle to adequately monetise risks for project appraisal. A small number of participants challenged the idea that existing practice fails to accommodate risk—they argued that flexibility of the Target Headroom mandated in the WRMP guidelines (Environment Agency, 2012) provides adequate scope for effective risk-cost trade-off.

In contrast to these views, several practitioners reported that a risk-based planning framework would provide additional rigour and a stronger foundation for decision making, but that the industry lacks the time, resources and technical capacity to undertake the extensive modelling assessments that underpin this approach. Some practitioners also identified business and regulatory risks associated with overhauling the planning methodology. For example, certain companies might be uncomfortable informing stakeholders that the previous analysis was “wrong” in the event that a new form of analysis recommends a radically different course of action. A number of company planners indicated that regulators are resistant to major change because they have invested a lot in current approach, which it is “tried and tested.” Some participants reported that the industry was collectively aware that existing practice is fallible, but that there has been a lack of demonstration of alternative approaches.

7.4 Discussion

7.4.1 The role of modelling assessments in planning

The interviews exposed an interesting tension relating to the adequacy and effectiveness of DO assessment (and corresponding supply-demand balance) for communicating the state of a WRZ. To illustrate, no participant was able to provide a logical line of reasoning for the use of a specific margin of Emergency Storage; many even discussed how this feature has been buried in the analysis and forgotten about. Yet many of the participants also lauded the supply-demand balance on the basis of "clear communication" of the state of the WRZ. This raises an intriguing question: how can the measure of supply be deemed "clear" if it relies on such a significant level of "buried" subjectivity? And why did some participants report that the supply-demand balance provides a "clear" form of communication whilst others reported the exact opposite? One suspects the answer lies in the use language rather than in a divergence in opinion. "Clear" is probably the wrong word; a more accurate term might be "screened." A deterministic measure of supply conceals complexity and thereby mitigates difficult questions. Few people would argue that "supply" should not meet demand or that a "deficit" should not be addressed, primarily because the information is presented in such a way that encourages the viewer to ignore the fact that "supply" is a complex and indeterministic concept. Conversely, there is plenty of scope for argument on whether a water company should aim for 75% or 95% certainty that it will maintain a restriction Level of Service of 1 in 20 years. Or should it aim for 95% certainty that the 1000-year drought will not cause catastrophic failure? Why not the 1 in 1500-year drought—and 98% certainty—just to be on the safe side? Would that be worth the £100 million investment? The scope for debate is endless, which is why one participant used the term "paralysis by analysis" to describe the risk of stifling a company's ability to build consensus and legitimise a course of action.

Indeed, most participants made a similar point in a less forthright way by suggesting that stakeholders would struggle to understand probabilistic performance metrics, or that presenting a supply-demand balance is the "only way to get an informed response." This angle is slightly (and probably unintentionally) disingenuous because it assumes that stakeholders understand, and are adequately informed by, a supply-demand balance. The reality is that the only person that understands a supply-demand balance is the person who makes the underlying assumptions necessary to define supply in a deterministic way. Even if outsiders rigorously inspect the publically available planning documents they may still fail to gain such understanding, as several companies either fail to report on particular assumptions or withhold technical documents that outline those assumptions. The point is not to question or castigate the motives of planners, but simply to highlight the paradoxical nature of statements that emphasise the need for "clear" communication through a deterministic metric of supply. Taking these statements for what they really mean, the need to build consensus and generate an impetus for action through the modelling output is evidently important to a large section of the planning community. If a move away from deterministic planning prevents or hinders a company's ability to invest in water resources infrastructure, then it also creates significant business and regulatory risk. Such risks have been largely ignored in the academic literature promoting fundamental change in the planning process and perhaps deserve more attention in future academic discussions on how to shape the planning process for WRMP19 and beyond.

7.4.2 Implications for existing planning practice

It appears that effective decision making and effective project implementation can be conflicting objectives. On one hand, planners seek to use modelling assessments to develop a clear picture of risk and to understand how different options might address areas of concern. Nearly all of the participants agreed that there is a case for more detailed examination of risks in the planning guideline in order to help achieve those aims. On the other hand, a large section of the planning community appears reluctant to communicate the state of their WRZs in uncertain terms; the modelling assessment must legitimise action as well as inform it. The interview responses provide some insights to inform a discussion of the potential issues associated with accommodating these conflicting goals in a deterministic planning framework.

The use of DO analysis and a supply-demand balance does not necessarily preclude more complex forms of decision making. On the contrary, many participants described additional components of decision making that can take place outside the DO assessment procedure. Knowledge of experienced practitioners, political considerations and separate analyses (e.g., vulnerability/resilience analysis) were cited as important sources of information that inform the ultimate investment decisions. Of course, accommodating these aspects will present a challenge in cases where perceived investment needs fail to match with the outputs of the prescribed modelling procedure. How will a company with a vulnerability/resilience problem, for instance, invest to deal with that issue if its WRZ is in a state of surplus? One can only assume that participants referring to the "flexibility" of existing practice meant that modelling assumptions can be adjusted in order to legitimise the desired course of action. This form of behaviour may seem illicit, but there are few alternatives for reconciling strict adherence to a modelling procedure for justifying decisions with the need to allow for important factors that cannot be easily captured through those modelling assessments.

Several participants acknowledged this issue and reported that the prescribed process primarily aims to set a consistent basis for asking questions—that is, companies are not strictly bound by the least-cost capacity expansion ethos and may reasonably deviate from it by, for instance, favouring a set of options that do not necessarily resemble "least cost" for addressing a supply-demand balance but provide some other benefit (such as dealing with vulnerability). Challenging this view, one participant reported that companies that have faced public inquiry for their WRMPs had suffered extensive scrutiny focused specifically on the parts of their plans that deviated from the prescribed process. Moreover, there are other regulatory reporting requirements, such as the Security of Supply Index (SOSI), that derive from the supply-demand balance, so one cannot assume that companies would always be able to deviate from it without implications. So whilst in theory a company may be able to accommodate alternative aspects in the existing process in a clear and transparent way, in practice there are business risks associated with this approach. Regulators should perhaps consider that the current setup and the potentially conflicting objectives of decision-making and project justification may incentivise manipulation of modelling assumptions.

There may be wider risks associated with the use of a supply-demand balance for communicating the state of a WRZ, particularly where those interpreting it are unaware of its limitations. Some participants considered this issue, envisioning potential problems where politicians get the wrong impression. For example, it is well acknowledged among the practitioner community that three consecutive dry winters is plausible hazard and a significant risk for southeast England. Yet faced with a supply-demand balance in a state of surplus, politicians may overlook this threat. Similarly, people outside the water resources planning community are unlikely to be aware that DO assessments cannot be fairly cross-compared across different companies. For example, The UK Infrastructure Transitions Research Consortium recently published a national assessment of water security using aggregated DO data (UK ITRC, 2014). If a consortium of infrastructure planning experts based at the country's most prestigious academic institutions failed to recognise what is concealed by the current planning metrics then what can be expected of Government policy makers and others viewing these assessments? As one participant noted, these issues of comparability could become increasingly important if and when companies begin to consider water trading.

7.4.3 Future research needs

The interviews exposed a number of industry concerns relating to the practical and theoretical basis for a risk-based planning framework informed by stochastic modelling assessments. Some of the issues may be resolved through improved communication between the academic and practitioner communities. For example, the view that synthetic drought scenarios generated by stochastic models are implausible would surely be refuted by most academics in this field. Those promoting the use of stochastic generators may therefore need to somehow clearly explain why their models produce credible information for use in water resources planning. The view that industry lacks the necessary software, data and models to undertake stochastic water resources modelling assessments could be challenged using industry studies reported in the academic literature (e.g., Harris *et al.*, 2013; Asefa *et al.*, 2014) or through modelling trials conducted as part of upcoming UKWIR research. Further method development and case demonstration will be required to give practitioners greater confidence that a risk-based planning framework can provide information that is useful for making and justifying investment decisions.

7.4.4 Limitations

The interview study captured the views of only fifteen practitioners. Moreover, the sample was confined to planners, consultants and regulators. Therefore the identified business risks may fail to adequately represent wider industry and company opinion. Company director interviews, for example, might have expressed a different set of views relating to the role of modelling assessments in building a consensus for action. Moreover, since the prospect of stochastic modelling is nascent in England and Wales, the scope for detailed discussion around emerging planning methods was severely restricted. There was considerable variation amongst the participants regarding level of understanding and perspective of what stochastic modelling assessments could achieve. The industry may offer a more balanced and considered perspective on stochastic modelling as the new methods find their way into UKWIR projects and national WRMP meeting agendas. A more detailed and expansive interview study would perhaps be of greater value in two or three years' time as the EA begins to define the guidelines for 2019.

7.5 Conclusions

1. The study found that most practitioners acknowledge a number of weaknesses in DO assessment methodology and believe there is a case for more detailed examination of risk in the planning process. However, few see a strong case for a fundamental shift toward a risk-based planning approach informed by stochastic modelling assessments.

2. The study exposed a number of business risks associated with shifting away from conventional planning methods. Most importantly, several participants indicated that modelling assessments play an important role in building a consensus for action, which is more easily achieved when the state of a WRZ is presented deterministically using the supply-demand balance.

3. The study found that most practitioners are sceptical about the practicality and utility of risk-based methods in water resources planning. Some hold reservations about the industry's capacity for undertaking stochastic modelling assessments. Others distrust the outputs that emerge from such analyses or envisage difficulties in using the outputs to inform investment decisions.

CHAPTER 8 DISCUSSION

The studies reported in this thesis are distinct from the contemporary UK-based literature cited in Chapter 2. Specifically, the case research has examined the water resources system modelling methods that underpin stochastic planning frameworks. So whilst a number of studies in this field have introduced existing approaches to dealing with uncertainty through a discussion of "headroom," and the "Economic Balancing of Supply and Demand" (e.g., Hall *et al.*, 2012a; Matrosov *et al.*, 2013a), the analyses reported in this thesis focus more explicitly on the impacts of modelling assumptions used to constrain and define supply-based (e.g., DO) and integrated (e.g., reliability) performance metrics. The study has sought from the outset to identify challenges associated with different planning methodologies; scant effort has gone into developing and promoting methods of decision-making under uncertainty. Moreover, those challenges were identified through dialogue with planners and other practitioners based on the various analyses and case studies that were undertaken. As a result, the study offers a fresh perspective on the problems of water resources system design, which should be of value to the industry ahead of discussions on whether and how to adopt new planning principles in the near future.

This following discussion draws on the five separate studies described in the previous chapters to highlight the main contributions relevant to the research questions that were introduced in Chapter 2, which are:

RQ1. What **practical methods** could planners deploy to begin exploring the use of stochastic modelling assessments for analysing water resources system performance?

RQ2. What **insights** might planners gain from running stochastic modelling assessments, and how do "stochastic yield analysis" and "risk analysis" compare in this regard?

RQ3. What **practical, conceptual and institutional challenges** might water companies face when attempting to plan using stochastic modelling assessments?

8.1 Practical methods

The study has developed four distinct practical approaches that planners could use to begin exploring integrated performance metrics and synthetic flows. These should not be considered methods of decision-making under uncertainty; rather they constitute simple and practical methods for implementing stochastic modelling assessments and then interpreting the results. All approaches were developed and demonstrated using existing company water resources system models and freely available supporting software and data.

The approach described in Chapter 3 applies the risk-based principles outlined by Brown and Baroang (2011) and Hall *et al.* (2012a) to define event probability curves for drought hazards denoted with reservoir storage triggers. The approach extends the understanding, relative to the risk-based approach described by Hall *et al.* (2012a), of the consequences of different events by mapping likelihood to hazard duration. The shape of the probability-duration curve informs the planner as to the relative likelihoods of more damaging long-duration events.

The approach described in Chapter 4 advances the above method by focussing on demand centres rather than reservoir triggers. The main innovation is in the visualisation of risk using a vulnerability surface that describes water shortage hazard according to the implicit shortfall duration and magnitude. The paper demonstrates how this form of performance measure could be used to identify previously unrecognised risks, which materialise when the system is stressed under alternative inflow sequences. Whilst the quantitative understanding of these risks is precarious (i.e., likelihood estimates will be conditioned by the input stochastic data), the process of stress testing the system and identifying potential vulnerabilities could provide useful insights to planners, particularly when examining resilience and comparing across different systems.

The method reported in Chapter 5 differs from the two prior studies in that a supply-based—rather than integrated—metric is assessed under alternative synthetic inflow replicates. The study therefore lies in the "stochastic yield analysis" quadrant of the matrix. The study demonstrates the use of synthetic flows for linking climate variables to yield, which facilitates a simple and effective climate impact assessment using multiple sources of climate projection data. Previous versions of this approach linked climate statistics to a performance measure based on "reliability" (e.g., Brown *et al.*, 2011, 2012; Hallegatte *et al.*, 2012; Moody and Brown, 2013; Ghile *et al.*, 2014). The version reported in this thesis is therefore more practical for planners that seek to retain supply-based performance measures whilst exploring the impacts of stochastic inflows and perhaps conducting a bottom-up climate impact assessment.

The case analysis presented in Chapter 6 develops and applies a stochastic method for correcting bias in water availability assessments. The method applies synthetic flow replicates and measures performance using an integrated metric that was devised specifically for the analysis. The modelling approach therefore lies in the "risk analysis" quadrant of the matrix. The paper demonstrates how planners might begin to use stochastic modelling assessments within the existing planning framework because the risk analysis is carried out separately with the results used to adjust the planning margins used in traditional capacity expansion planning.

8.2 Insights: "stochastic yield analysis" versus "risk analysis"

If the goals of stochastic exploratory analysis of a water resources system are to transparently expose vulnerabilities, then "stochastic yield analysis" is a poor substitute for "risk analysis." The reason is that yield—or Deployable Output—does not distinguish the nature of failure that a system would experience under given drought conditions—as demonstrated in Chapter 4. So even when the yield is re-examined under a range of stochastic inflow sequences, the planner remains uninformed as to the severity of shortfall that customers could experience under those droughts. The point was raised by reservoir storage theorists decades ago: "The weakness of the [non-failure] approach becomes apparent if we realize that it does not distinguish between a case where a given storage would result in a failure in water supply lasting, say, a year with supply falling to 10% of the target and a case where a failure would last a couple of days during which the supply would have been reduced to 95% of the target" (Klemes, 1981). Similarly, the analysis of water resources system yield masks over these details and prevents an analysis of risk because the tangible consequences of severe drought and system failure are not adequately expressed. The interview study in Chapter 7 established that this level of detail is important for planners and most participants acknowledged that the existing process fails to adequately capture vulnerability and resilience issues.

The impact of the "blind spot" imposed by yield analysis on investment decision making can be brought to light with a very simple example. Consider a simple single reservoir system providing water to a single demand centre. Suppose that system is in supply-demand deficit of 100 ML/d and that the incumbent company is considering two possible options for addressing the deficit. The first option involves raising the reservoir weir to increase storage, and costs £90m. The second option involves drilling a borehole to make available a plentiful groundwater source, which would maintain a given level of supply under drought conditions and would cost £100m (we can ignore the operational costs for simplicity). Suppose both schemes provide exactly 110 ML/d additional DO at those costs. According to the least cost capacity expansion paradigm, the optimal investment choice is the reservoir extension, offering equal DO to the borehole but at less cost. However, this analysis discounts the fact that, under conditions of reservoir failure, the borehole option would sustain a given level of supply. Failure under the reservoir extension option would be catastrophic (i.e., 100% demand shortfall), since the demand centre would have no alternative sources from which to draft water.

In the real-world, this vulnerability blind spot may be less palpable than in the example above. Nonetheless, it stands to cause oversight in system design and will manifest in inefficient plans. The oversight is not necessarily confined to supply-based performance metrics; it is imposed on any analysis (supply-based or integrated) that defines the performance of a water resources system based on a limited number of reliability thresholds, which would include the studies described by Hall and Borgomeo (2013); Matrosov *et al.* (2013a), Kasprzyk *et al.* (2013) and others.

Though also based on thresholds (reservoir triggers), the study reported in Chapter 3 provides some additional insights to overcome the vulnerability blind spot because it extends the description of failure by also capturing the failure duration. However, the study presented in Chapter 4, whilst based on limited-length input data, provides even clearer insights of the state of the system and the performance of alternative strategies because the failure data is disaggregated into numerous impact types, ranging from short duration disruptions through to long duration catastrophic failures. The difficult unanswered question is: how can an understanding of resilience and vulnerability, developed through the type of exploratory analysis described in Chapter 4, be used to inform investment decision making? This prospect is discussed in the next section.

Despite this challenge, there may still be some useful insights to be gained from a "stochastic yield analysis." The study reported in Chapter 5 demonstrates how such an approach could be used to test the sensitivity of system performance to climate statistics, thereby facilitating a "bottom-up climate impact assessment." Companies in England and Wales might adopt this approach to determine whether their systems are sensitive to the climate variables that are adequately modelled by GCMs. This type of analysis could be used to pre-determine the worth of, and need for, a top down impact assessment based on the UKCP09 projections.

"Risk analysis" was also found to be potentially useful in comparative studies—i.e., comparing risks across alternative water resources systems and across alternative options. The study reported in Chapter 6 demonstrates the use of a custom-made metric based on reservoir drawdown behaviour during drought, which be compared across systems to bias-correct "Emergency Storage." The insights would be useful for companies that operate significantly different WRZs and that define DO using the drought of record (such as United Utilities).

8.3 Practical, conceptual and institutional challenges

8.3.1 Practicalities

The four case studies presented in Chapters 3, 4, 5 and 6 demonstrate that stochastic yield analyses and risk analysis of water resources systems can be executed using existing company models, freely available data, open-source software and some additional code that could be written by any well trained hydrologist. The methods employed impose lengthy run times compared to traditional yield analysis. However, given the levels of investment that hinge on planning decisions, some additional computing effort should not deter the industry from changing practice if the associated methods can be shown to improve rigour. There are simple technologies and techniques that companies could implement to drastically speed up computation without overhauling existing software packages and resource system models. High performance computing technologies, such as distributed computing and Graphics Processing Units (GPUs), could reduce run times by orders of magnitude and have recently been demonstrated in a practical water resources planning context (Bryan, 2013; Asefa *et al.*, 2014). So the argument that stochastic modelling is impractical is a weak one. Instead, the arguments that require closer examination are those that question the utility of planning frameworks based on stochastic modelling assessments. The interviews and case studies identified some conceptual challenges that seem to invalidate some of the possible benefits that are argued in the literature, particularly those relating to transparent planning.

8.3.2 Conceptual challenges

Subjectivity in yield analysis

The issue of subjectivity in a yield-based capacity expansion planning framework appears several times in the articles presented above. All participants in the interview study acknowledged that, when planning using the critical drought in a yield-based framework, there is a need for an arbitrary reserve storage margin to allow for the possibility of a more extreme event. But some interview participants reported that this subjectivity could be evaded in situations where yield is constrained using desired levels of service based on frequency of non-essential use bans (based on a storage trigger position). This supposition needs to be addressed here because it implies that the issue of subjectivity, which impacts the transparency of the outputs, might easily be resolved within a traditional planning framework—that is, without risk analysis.

Consider two hypothetical systems—Zone A and Zone B. They are identical in nearly every way. Both systems must meet the same demand and both are single reservoirs fed directly by the same single catchment inflow, for which there is a 100 year record. The failure criterion that constrains yield in both systems is a maximum 1 in 25 year level of service condition for non-essential use bans. The control curve that initiates the non-essential use ban is defined by the storage behaviour under the historical conditions and assuming existing demands—by definition it must lie at the same storage depletion level in both systems. The yield is therefore equal to the maximum supply that the system can sustain without drawing the storage below the control curve more than four times over the 100-year simulation of the historical inflows. When the supply demand balance is presented to the customers they are informed that the supply in the diagram represents the supply that ensures their desired level of service is met (of course, supply is identical in both systems). The assessment identifies a looming deficit in both systems that manifests as the demand increases within the planning horizon. Has all subjectivity been removed such that both sets of customers are presented with a clear view of the state of the system and a comprehensive case for capacity expansion to meet their desired level of service?

Perhaps surprisingly, the answer is no.

The final crucial piece of information is that the total effective storage capacity in Zone A is twice that of Zone B, as depicted in Figure 8-1. Clearly Zone A must be significantly more robust than Zone B—and yet the yield in both systems is identical and the case for investment in both systems is also seemingly identical. But the analysis for both systems was carried out in the same systematic fashion. So how has this situation arisen? The answer lies in the definition of the restriction trigger position. The use of historical reservoir behaviour to define the restriction trigger position creates the illusion that the trigger has been positioned objectively. The problem, however, is that the trigger positioning bears no relation to the available storage and the risk of total storage depletion and associated catastrophic failure. The customers of Zone B tolerate a much greater risk of catastrophic failure than the customers of Zone A, although they would be unable to interpret this fact from the supply demand balances presented. This is what is concealed if the yield is defined using the level of service constraint rather than the critical drought.

Of course, there are several ways in which to position restriction triggers. The analyst might look to position the trigger to reflect the relative risks in either system. But without a clear and quantified understanding of the relative consequences of restriction versus catastrophic failure it will be impossible to objectively position these triggers in a way that provides transparency in a single supply-demand balance. In practice the trigger position relies as much on human psychology, politics and engineering judgement as it does on hydrology and science (Lambert, 1988). So the supply component of the supply-demand balance will depict the maximum supply that can be maintained whilst meeting the desired level of service, *assuming an arbitrary restriction trigger position*. Theoretically, the analyst could drive the yield up or down by adjusting the restriction trigger position, similarly to how an analyst might alter the reserve storage volume to drive a critical-failure yield value up or down. In practice this behaviour is unlikely because control curve positions are written into operational agreements just as required emergency storage volumes are defined (albeit loosely) in planning guidelines. But just because something is written in a regulatory document does not preclude it from being subjective and therefore potentially misleading when presented in a deterministic supply-demand balance.

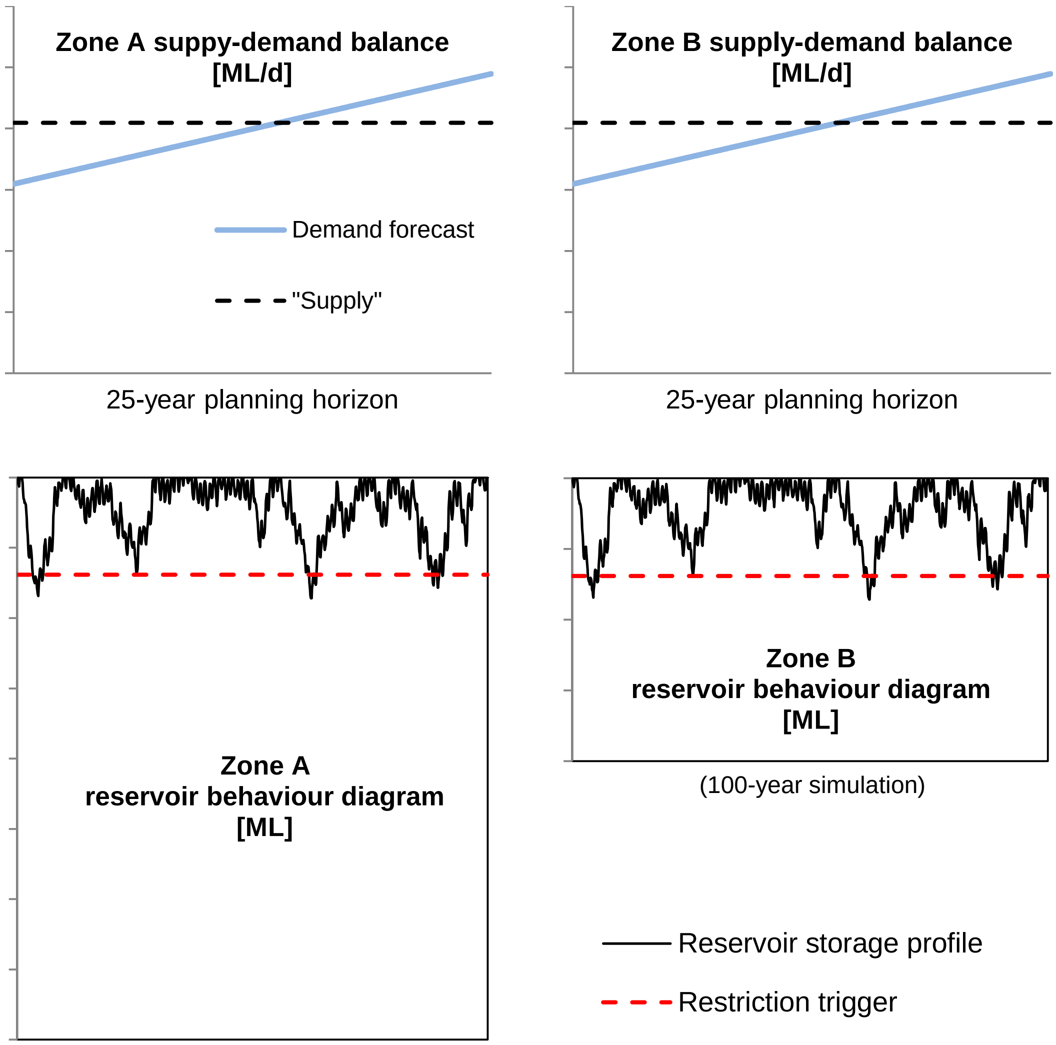


Figure 8-1 Two hypothetical water resources systems producing identical supply-demand balances despite difference in storage.

Subjectivity in risk analysis

In general, arbitrary assumptions are an accepted and necessary part of yield-based planning. This is not necessarily a limitation if the aim of the plan is to define a baseline for scenario and option testing, particularly when there is no need for consistency across different systems and companies. But if the planner aims to transparently trade off costs of investment against water shortage risks then the subjectivity becomes problematic. So to what extent do the planning approaches based on risk analysis overcome subjectivity?

Consider the probabilistic approach described in Chapter 3. Hall and Borgomeo (2013) argue that "transparently implemented risk analysis provides a mechanism for exposing the implications of uncertainty for outcomes that people value." They provide a case example with results highlighting the performance of different investment options in terms of probability of failing to meet a target level of service for a temporary use ban. But this, on its own, is an inadequate performance measure if the aim is to transparently expose outcomes that people value. The issue is similar to that presented in the hypothetical example above (Figure 8-1). If the performance measure relies solely on the restriction risk then the (technically) optimal option for system improvement will be to simply lower the restriction control curve, which incurs no financial investment costs and reduces the probability of failing to meet the desired service level. Of course, in the real world, the water company cannot simply lower the control curve; the longer a company waits to impose restrictions the more likely the system will be to run out of water. But the point is that the proposed risk analysis does not capture the risk of running out of water. Nor does it capture the resilience and vulnerability of the system under total failure conditions. Transparency may only be achieved if all types of risk are presented and then traded off against each other using control curve level adjustment alongside all other proposed measures for system improvement. This will be impossible if companies wish to present a plan with a catastrophic failure probability of "never" and will be extremely challenging if customers and stakeholders are to define the tolerable probabilities of suffering catastrophic failures. The analysis that sidesteps extreme risks also stands to discount the value of investment options that are effective in tackling those risks through resilience-enhancing options, such as climate-independent sources of supply.

The same general argument can be applied to any planning framework that uses integrated performance metrics based on fixed storage level triggers. Consider the analysis described by Matrosov *et al.* (2013a), which measures "engineering robustness" using a restriction trigger. It is evident from a basic understanding of the equations presented that a different set of "robust" decisions would have been generated by the analysis if the restriction control curve had been adjusted or if an alternative storage level (e.g., reservoir failure) had been used to measure the performance. Given this challenge, can the proposed method really serve a planner any better than the basic process of simulating the system under synthetic flows, looking for vulnerabilities and then defining a plan using some rudimentary engineering judgement? It is hard to see a good reason why a method that produces a crisp solution with elusive subjectivity should be more useful to a planner than a method that acknowledges its basic flaws.

A number of studies have looked to develop Pareto portfolios of investment solutions using multi-objective optimisation. Taking a similar line of argument to that offered above, Mortazavi *et al.* (2012) reviewed the state of the art in this field and then demonstrated how the common failure to consider catastrophic failure risk and the full decision space (i.e., control curve adjustment) can lead to inferior solutions. The authors then presented their own approach that incorporates these risks, but acknowledged that their Pareto portfolios would be sensitive to the length of synthetic record and return period of the most extreme drought in the analysis. The question for a planner that might have to use these analyses is: why select an investment plan from a Pareto frontier that assumes the 10,000-year drought when the analysis would return a different Pareto frontier (with different solutions) assuming a 5,000-year drought? It seems that no matter what steps are taken, the need for the arbitrary assumption re-emerges in a different place. The point is not to suggest that this class of method should be avoided or written off, but rather to highlight the limitations, which, if overlooked, could lead the decision maker astray.

Approaches that combine robust decision making with Pareto optimisation have been developed and proposed more recently (e.g., Kasprzyk *et al.*, 2013). These studies have pushed computational and theoretical boundaries in this field to generate candidate solutions that dominate across several objective functions and simultaneously achieve adequate performance under extreme drought conditions. Data visualisation techniques, which can present up to seven dimensions on a single plot, expose some of the many trade-offs planners might need to consider. Perhaps this route of method development will ultimately yield an approach that provides demonstrably transparent and valuable insights for real-world decision making. But as yet there remain fundamental issues with the approach. For example, Kasprzyk *et al.* (2013) use a measure of reliability ("critical reliability") as a boundary constraint, such that a candidate solution must ensure at least 60% of demand is met under critical drought conditions (i.e., a maximum shortfall of 40%) across at least 99% of the modelled scenarios. One must expect that the resulting Pareto portfolios would be dominated by solutions adopting more dependable supply sources if the maximum shortfall constraint were ramped up to, say, 70% demand. Yet the authors provide no reasoning for their particular boundary constraint, nor do they address how it might impact the solutions and visual outputs. Ironically, decision makers are expected to draw from these analyses valuable insights that overcome their "cognitive myopia," which supposedly arises when they "inadvertently ignore aspects of the problem... such as important decision alternatives or key planning objectives." Having closely examined these assumptions—particularly in Chapter 4, which unpacks the full range of water shortage risks and then struggles to aggregate them in any useful way—this thesis finds that any attempt to rigorously use the results of vulnerability assessments to inform investment strategy may conceal valuable detail. In light of this challenge, a worthwhile question would be whether a demonstrably robust solution (or robust Pareto frontier) can plausibly exist without a rigorous method for aggregating the full range of drought-related shortfall risks, from short-lived minor demand shortfalls (implemented through restrictions) through to mega-droughts that cause societal, economic and environmental catastrophe. To ignore the full range of risks is to conceal from the end user a full and transparent description of the state of the system and its inherent trade-offs.

The industry has been able to avoid these difficult questions with planning heuristics: reserve storage margins, control curve positions, conservative assumptions of available groundwater levels, ideal flows apportioned to the environment at different stages of drought, and so on. These assumptions are never completely objective and rarely inconsequential to the decision. They are a fundamental and necessary foundation for developing a baseline for use in traditional planning. But if one aims to transform the nature of planning—such that the process generates decisions that are demonstrably more transparent, more robust, more effective, science-based in some way—then the premise and reasoning for planning heuristics no longer holds. Indeed, the need to remove elements of subjectivity in a transparent risk analysis appears to be an accepted point among those that have critiqued the use of Headroom (e.g., Hall *et al.*, 2012a). The same argument must also apply to the many other less obvious subjective judgements that influence the assessment.

Real consequences

The problem of modelling "real consequences" is mentioned in Chapter 3 and was brought up by a number of practitioners in the interviews described in Chapter 7. What does this mean and why is it a fundamental challenge for transparently exposing vulnerabilities using exploratory modelling? Invariably, a water resources risk analysis is informed by simulated events that occur during drought—restrictions, supply shortfalls, reservoir failures, etc. The decision making that follows generally aims to weigh up the worth of investments designed to reduce the likelihood of those events occurring. If this is to be done seriously then the analyst must consider major hazards, including catastrophic system failure. The problem is that companies may be unable to adequately model system response under these conditions.

There are a number of reasons this issue might manifest. Perhaps most obviously, a state of civil emergency, hastened by the threat of a severe water shortage, could generate sufficient political impetus for unconventional actions. Environmental thresholds, licence conditions and compensation flow requirements—represented as hard constraints in system models—may no longer apply when the alternative is rota cuts and standpipes. It is difficult to imagine that United Utilities, for example, would impose rota cuts on seven million people in northwest England whilst usable water could be drafted from the "hands-off" reaches of the Lake District. Or perhaps in other situations the compensation flow requirements would be relaxed and abstractions made available from environmentally sensitive rivers. The available options would depend not only on the physical geography of the system and catchments, but also on the political climate and the perceived urgency of the situation. The response would be impossible to predict to any reasonable degree of accuracy.

Another issue is that many systems have never experienced the extreme events that might cause catastrophic failure and so the analyst might lack the information necessary to predict how much usable water would be physically available under those circumstances. For example, many water companies will be unaware of how the quality of water in their reservoirs will be affected once the "dead water" levels are reached. Similarly, some companies have been unable to quantify the flows of water that would be available from groundwater sources at unprecedented low levels (as one practitioner revealed during the interview study). Moreover, modellers and analysts that have not experienced extreme drought situations may overlook real options that could be made available under extreme circumstances, such as using previously abandoned sources or tankering water. Whilst these actions could theoretically be captured in company drought plans (and then in the models), a real drought situation rarely plays out in the same way (Watts *et al*., 2012).

The reason all of these aspects are important is that if the modelled understanding of when events, such as rota cuts, would actually occur is skewed then any assessment of risk would also be skewed. In certain cases the modelled likelihood estimates for these impacts could be miscalculated by orders of magnitude. For example, if non-modelled options for extending supply by, say, breaching licence conditions or finding water from an alternative source, provide an additional two months of supply, then that may be enough to avoid failure prior to winter months, during which system would be more likely to be recharged. So a 1 in 1,000 year likelihood estimate for implementing rota cuts may become a 1 in 10,000 year likelihood estimate once all the other relevant factors are taken into consideration. This would be system dependent, but given the potential for widely misplaced estimates it represents a major challenge for a transparent risk analysis of any kind. The issue is not necessarily restricted to planning methods that employ probabilistic metrics because these assumptions are propagated into the decision analyses and visual outputs in any given decision-making method that relies on risk analysis.

8.3.3 Institutional challenges

Daniel Loucks stated that "planning and managing involves not only decision making, but also developing among all interested and influential individuals an understanding and consensus that legitimizes the decisions and enhances their successful implementation" (Loucks, 1992). The interview study presented in Chapter 7 demonstrates the veracity of this statement in the context of England and Wales. The most important institutional challenge identified was the widespread concern that exposing difficult risk-cost trade-offs might cause "paralysis by analysis"—endless, non-productive debate that prevents or hinders a company's ability to build consensus for action.

The contemporary literature on water resources planning methodology offers very limited discussion on this issue, which directly challenges the view that transparency is a desirable feature of planning metrics. The paradox here is that a transparent view of the state of the system, which is supposed to be helpful for planners in deciding how to act, may simultaneously hinder the prospects for legitimising planning decisions. The reason this paradox occurs is very simple: making the assessment transparent involves stripping away the precarious and arbitrary assumptions so that what is left is a bare-bones portrayal of the underlying challenge. The discipline of water resources planning becomes exposed for what it actually is: an intractable design problem, plagued by severe and irreducible uncertainties, for which there is no logical analytical solution. The only way to make such a problem manageable—to allow the analysis to work its decision-making magic and then perhaps fabricate some legitimacy—is to paper over the cracks by reinstating the precarious assumptions, subjectively defined margins, concrete boundary constraints... in other words, all the things that obscured the problem from transparency in the first place!

The study reported in Chapter 7 identifies a number of other possible institutional challenges associated with change in planning methodology toward a stochastic approach. Some planners fear that new (and possibly expensive) decisions recommended by a new form of analysis might raise suspicion and mistrust among customers. Some believed that customers would naturally distrust any form of analysis based on synthetic droughts, although the plans recently submitted by Southern Water will test the validity of this concern (Southern Water, 2013). A possible additional problem might be that the entire industry—including its many consultants—is used to and understands a particular form of planning. Therefore, one might expect substantial transition costs as the industry adapts to a new way of determining and communicating water shortage risks.

CHAPTER 9 CONCLUSIONS

9.1 "Is the pain worth the gain?"

The above question was pondered out loud by a planner during interview. It succinctly captures two common assumptions about stochastic modelling: first, that there is significant pain; second, that there is significant gain. The study reported in Chapter 7 exposed a general belief, particularly amongst regulators, that stochastic modelling would necessarily (and vastly) improve planning rigour, but would also be impractical, as companies lack the necessary skills, resources, models and data to execute the analyses. This thesis finds these beliefs to be at best exaggerated, at worst misconceived. New and valuable insights are by no means guaranteed by stochastic modelling assessments. Where new insights are available, the industry may lack the tools to use them effectively to inform investment planning decisions.

The thesis has introduced a novel matrix to define water resources modelling methods using two dimensions: measure of performance and treatment of hydrological uncertainty. The resulting matrix separates "traditional planning" from two stochastic modelling approaches, which have been termed "stochastic yield analysis" and "risk analysis." Every stochastic modelling study referenced in this thesis—irrespective of whether the overarching planning framework is based on multi-objective optimisation, robust decision making, real options analysis, or any other decision-making approach—falls into one of these two categories.

"Stochastic yield analysis" could be useful for planners that wish to express the impact of hydrological uncertainty on DO assessments. It could also be used to conduct a bottom-up climate impact assessment. The study presented in Chapter 5 developed the first instance of such an assessment and demonstrated its applicability to large and complex water resources systems. Thus, assessments based on stochastic yield analysis could form replacements or complementary additions to some of the existing features of the WRMP process. However, yield analysis requires subjectively defined planning margins and a binary definition of system failure. A planning framework based on stochastic yield analysis inherits these simplifying assumptions and then adds more. The resulting analysis adds little to the understanding of the system that could enable a planner to generate investment plans that are demonstrably more robust than the status quo. The process could be used to legitimise investments to improve system reliability—but then so could the simple arbitrary addition of 10% onto an arbitrary planning margin. The former is no more rigorous than the latter from a technical point of view.

The thesis finds that water resources planners in England and Wales may be able to improve their understanding of water shortage risk—relative to the level of understanding achieved using Deployable Output analysis—by extending the resource system analysis through the use of synthetic flows and integrated performance metrics. This process is termed “risk analysis” in the matrix. Risk analysis can characterise the nature of failure—that is, the magnitude and duration of shortfalls—that population centres may experience under particular drought scenarios. By exposing these risks, the process could nurture a more transparent description of the vulnerability of a water resources system. The thesis finds that risk analysis can be executed using current industry software, models and data. This conclusion challenges the widely-held view, exposed during the interview study, that the companies are ill-equipped to conduct planning frameworks based on stochastic modelling assessments. The case demonstrations—to be published in international journals—may give confidence to company planners, modellers and analysts considering a closer examination of water shortage risk in comparative studies, drought planning and possibly as part of routine efforts to understand the impacts of changes to licence constraints, compensatory release requirements and so forth. They might employ the well-established metrics of “failure duration” and “water deficit” (Klemes, 1981), and could visualise the risks using something akin to the vulnerability surfaces developed in Chapter 4. Moreover, by understanding the stochastic nature of system performance—and having removed the blind spots imposed by DO analysis—the risk analysis could be useful for targeting a given level of investment within and across WRZs, as demonstrated in Chapter 6.

But a more transparent view of the system vulnerabilities does not necessarily translate to a well-informed investment plan. Transparency implies exposing the severe uncertainties that affect the analysis. One well-established cause of severe uncertainty is the hydrological scenarios—neither non-stationary stochastic models nor GCMs provide data that can inform reliable future drought frequency estimates. Compounding this issue, the thesis identifies a problem of “real consequences”; companies may be unable to adequately model the system response under severe conditions, invalidating the modelled thresholds that are assumed to correlate with impacts. Thus, whilst a stochastic modelling exercise can reveal vulnerabilities, there exists no process that can determine how likely these will be to manifest in future. It appears that any attempt to progress from transparent descriptions of vulnerability toward some form of strategy recommendation requires arbitrary assumptions. Such assumptions violate the underlying rigour and transparency of the analysis and could potentially cause, rather than cure, cognitive myopia.

Of course, none of this helps the planner, who must reach a decision regardless. The existing planning process provides some simple crutches that help recommend and legitimise action despite the wicked nature of the problem. The interview study in Chapter 7 finds that many planners value the existing process for this very reason. The existing process is also adaptive: one thing that is guaranteed to create supply-demand deficit—and then impetus for investment—is more severe droughts than those experienced in the last century. The risk, of course, is that a future drought is so severe that it depletes stored water and causes catastrophic failure. Useful future research might focus on attempting to improve the modelled understanding of the real impacts on customers and the environment by incorporating contingency measures into reliability and resilience studies.

9.2 Veracity of conclusions and recommendations for further work

The veracity of the conclusions may be jeopardised to some extent by the limited number of modelling trials and interviews conducted. The final paragraphs of the thesis offer up some possible limitations as well as recommendations for future research. Table 9.1 then summarises the main contributions.

The thesis finds that stochastic modelling of water resources systems should be technically feasible from a company point of view. This finding requires further qualification. The thesis does not mean to imply that companies could easily and effortlessly adopt stochastic modelling into their regular planning activities. The studies conducted through this research project were one-off attempts to run stochastic simulations of water resources systems using existing models. Company modellers are required to run simulations and capture system performance data on a daily basis for various different practical purposes other than developing their WRMPs. It is difficult to imagine—at least at this stage—that companies could substitute DO analysis for stochastic assessments in every modelling study they undertake. Therefore, whilst the industry is well-equipped to begin exploring the use of stochastic modelling as part of the planning process, there would significant challenges for a full transition away from "traditional yield analysis" in day-to-day modelling activities. Future industry research might examine the set-up costs for high-performance computing capacity that would allow for significantly faster simulations using the existing modelling software.

The modelling trials were based on a small number of water resources systems—all predominantly fed by surface water resources, such as rivers and reservoirs. The generality of the conclusions, particularly those relating to technical execution challenges, is therefore somewhat limited. Aquator, for example, does not handle the detailed groundwater modelling that might be necessary in the chalk dominated systems of southeast England. A more extensive, generalised analysis using a broader range of software platforms and models would be required to strengthen the conclusions. Additionally, the interview study captured the views of only fifteen practitioners. The sample was confined to planners, consultants and regulators. Therefore the identified business risks may fail to adequately represent wider industry and company opinion. Company director interviews, for example, might have expressed a different set of views relating to the role of modelling assessments in building a consensus for action. Moreover, since the prospect of stochastic modelling is nascent in England and Wales, the scope for detailed discussion around emerging planning methods was severely restricted. The industry may offer a more balanced and considered perspective on stochastic modelling as the new methods find their way into UKWIR projects and national WRMP meeting agendas. A more detailed and expansive interview study would perhaps be of greater value in two or three years' time as the industry regulators begin to define the guidelines for 2019.

Finally, the thesis has challenged the utility of emerging water resources planning methodologies based on stochastic assessments. These challenges are technical, focusing on modelling assumptions and input data. The critique may overlook the importance of insights that analysts can acquire through the actual process of conducting complex modelling assessments; constructing a model, collecting data, devising performance metrics, running simulations, figuring out causal relationships, etc., may endow the analyst with a level of understanding of the system that pervades high level decision makers (Lund, 2012). Such insights will be of real practical value in planning discussions and should not be discounted.

Table 9.1 Summary of contributions

Description	**RQ**	**Chapter(s)**
The research shows that stochastic risk analyses can be executed using bespoke industry water resources modelling software and freely available data, including for large multi-basin systems.	3 – practicalities	3, 4, 5, 6
The research identifies conceptual challenges for informing water resources planning decisions using risk analysis, including: subjectivity in performance metrics; lack of methods for trading off different types of hazard (duration/magnitude); limited understanding of the "real" consequences of severe drought.	3 – conceptual challenges	3, 4, 5, 6, 7, 8
The research identifies business risks associated with a change in regulated planning practice, including hampering a company's ability to build consensus for an investment plan.	3 – institutional challenges	7
The research has developed "event probability profiles"—a simple diagrammatic representation of risk in terms of hazard duration and likelihood.	1 – method development	3
The research demonstrates how a stochastic analysis of a weakly-interlinked system can identify previously unrecognised vulnerabilities and thereby alter understandings of cost-effectiveness of different investment options.	2 – useful insights	4
The research has developed a novel vulnerability matrix for characterising demand shortfall risks identified through stochastic assessment in terms of magnitude, duration and likelihood.	1 – method development	4
The research demonstrates that a decision scaling climate impact assessment methodology can be applied to large and complex water resources systems.	3 – practicalities	5
The research has developed a yield-based decision scaling approach, which enables simple incorporation of demand forecast uncertainty in a bottom-up climate impact assessment.	1 – method development	5
The research demonstrates that traditional yield-based assessments that rely on a reserve storage assumption cannot be fairly cross-compared.	N/A – additional contribution	6
The research has produced a simple stochastic method for addressing reserve storage bias in critical drought ("no failure") yield assessments.	1 – method development	6

REFERENCES

Adamson, D., Mallawaarachchi, T. and Quiggin, J. (2009), Declining inflows and more frequent droughts in the Murray–Darling Basin: climate change, impacts and adaptation, *Australian Journal of Agricultural and Resource Economics*, 53, 345–366. doi: 10.1111/j.1467-8489.2009.00451.x.

Ajami, N.K., Hornberger, G.M. and Sunding, D.L. (2008) Sustainable water resource management under hydrological uncertainty, *Water Resources Research*, Vol. 44, W11406, doi:10.1029/2007WR006736.

Allen, M.R. and Ingram, W.J. (2002), Constraints on the future changes in climate and the hydrological cycle, *Nature*, 419, 224-232.

Anagnostopoulos, G.G., D. Koutsoyiannis , A. Christofides , A. Efstratiadis & N. Mamassis (2010) A comparison of local and aggregated climate model outputs with observed data, *Hydrological Sciences Journal*, 55:7, 1094-1110.

Archfield, S. and Vogel, R. (2005) Reliability of reservoir firm yield determined from the Historical drought of record, Impacts of Global Climate Change: pp. 1-8.

Arnell, N. W. (2011), Incorporating climate change into water resources planning in England and Wales, *Journal of the American Water Resources Association*, 47: 541–549. doi: 10.1111/j.1752-1688.2011.00548.x

Asefa, T., Clayton, J., Adams, A., Anderson, D. (2014) Performance evaluation of a water resources system under varying climatic conditions: Reliability, Resilience, Vulnerability and beyond, *Journal of Hydrology*, 508, 53-65, doi: 10.1016/j.jhydrol.2013.10.043.

Barma and Varley (2012) Hydrological modelling practices for estimating low flows – guidelines, Low flows report series, National Water Commission, Canberra, Australia.

Barsugli, J. J., Vogel, J. M., Kaatz, L., Smith, J. B., Waage, M. and Anderson, C. J. (2012), Two faces of uncertainty: Climate science and water utility planning methods, *J. Water Resour. Plann. Manage.*, 138(5), 389-395. doi: 10.1061/(ASCE)WR.1943-5452.0000188.

Barnes, F. B. (1954) Storage required for a city water supply, *Journal of the Institution of Engineers.*, Australia, 26, 198.

Beard, L. R. (1965) Methods for determination of safe yield and compensation water from storage reservoirs, US Army Corps of Engineers – Technical Paper Series (TP3), Davis, California.

Ben-Haim, Y. (2006), Info-gap decision theory: Decisions under severe uncertainty, 2nd Edition, Elsevier, Oxford, UK.

Bertsekas, D.P. and Tseng, P. (1994) RELAX-IV: A faster version of the RELAX code for solving minimum cost flow problems, *Completion Rep. Under NSF Grant CCR-9103804*, Department of Electrical Engineering and Computer Science, MIT, Cambridge, Mass.

Beven, K. (2011) I believe in climate change but how precautionary do we need to be in planning for the future? *Hydrological Processes*, 25 (9), 1517-1520, doi: 10.1002/hyp.7939.

Blöschl, G. and Montanari, A. (2010), Climate change impacts—throwing the dice?. *Hydrological Processes*, 24, 374–381. doi: 10.1002/hyp.7574.

Borison, A., Hamm, G., Farrier, S. and Swier, G. (2008), Real options and urban water resource planning in Australia, *WSAA Occasional Paper No. 20 - ISBN 1 920760 30 X*, Water Services Association of Australia (WSAA), Melbourne, Australia.

Boughton, W. (2004). The Australian water balance model. *Environmental Modelling & Software*, 19(10), 943-956, doi: 10.1016/j.envsoft.2003.10.007.

Brown, C and Baroang, K. M. (2011), Risk assessment, risk management, and communication: Methods for climate variability and change, in *Treatise of Water Science*, vol. 1, edited by P. Wilderer, pp. 189-199, Elsevier, New York.

Brown, C., Ghile, Y., Laverty, M. and Li, K. (2012), Decision scaling: linking bottom-up vulnerability analysis with climate projections in the water sector, *Water Resources Research*, 48, W09537, doi:10.1029/2011WR011212.

Brown, C., Meeks, R., Ghile, Y., and Hunu, K. (2013) Is water security necessary? An empirical analysis of the effects of climate hazards on national-level economic growth, *Philosophical Transactions of the Royal Society A*, 317: 20120416, doi: 10.1098/rsta.2012.0416.

Brown, C., Werick, W., Leger, W. and Fay, D. (2011), A decision-analytic approach to managing climate risks: Application to the Upper Great Lakes. *Journal of the Americal Water Resoures Association*, 47: 524–534. doi: 10.1111/j.1752-1688.2011.00552.x.

Brown, C. and Wilby, R. (2012), An alternative approach to assessing climate risks, *Eos Trans. AGU*, 93(41), 401-402. doi: 10.1029/2012EO410001.

Bryan (2013) High-performance computing trials for the integrated assessment and modelling of social-ecological systems, *Environmental Modelling and Software*, 39, 295-303. doi: 10.1016/j.envsoft.2012.02.006.

Catawba-Wateree Water Management Group (2014) Catawba-Wateree River Basin water supply master plan, Prepared by HDR and McKim & Creed, Charlotte.

Chambers, V.K., Creasey, J.D., Glennie, E.B., Kowalski, M. and Marshallsay, D. (2005) Increasing the value of domestic water use data for demand management, WRc plc, Swindon.

Chen, C., Huang, C.H., Li, Y.P., Zhou, Y. (2013) A robust risk analysis method for water resources allocation under uncertainty, *Stochastic Environmental Research and Risk Assessment*, 27, 713-723, doi: 10.1007/s00477-012-0634-5.

City of San Diego Public Utilities Department (2013) 2012 long-range water resources plan, San Diego, California.

Davies, A. and Daykin, S. (2011) Review of water resources management planning process: final report, Policy projects for CGL, DfT, DECC and Defra, PB13653, IHRP, London, UK.

de Neufville, R. (2004) Uncertainty management for engineering systems planning and design, Engineering systems monograph, MIT Press, Cambridge, MA.

Daron, J.D. and Stainforth, D.A (2013), On predicting climate under climate change, *Environmental Research Letters*, 8(3), doi:10.1088/1748-9326/8/3/034021.

Department of the Environment (1996) Water resources and supply: Agenda for action, ISBN: 9780117532991, DoE, London, UK.

Dessai, S. and Hulme, M. (2007) Assessing the robustness of adaptation decisions to climate change uncertainties: A case study on water resources management in the East of England, *Global Environmental Change*, 17, 59-72, doi: 10.1016/j.gloenvcha.2006.11.005.

Dessai, S., Hulme, M., Lempert, R. and Pielke, R. Jr. (2009), Climate prediction: a limit to adaptation? in *Adapting to climate change: Thresholds, values, governance*, edited by W. Neil Adger, Irene Lorenzoni and Karen L. O'Brien, pp. 64-78, Cambridge University Press, Cambridge, UK.

Dessai, S., Browne, A. and Harou, J. J. (2013), Introduction to the Special Issue on Adaptation and Resilience of Water Systems to an Uncertain Changing Climate, *Water Resources Management*, 27(4), 943-948. doi: 10.1007/s11269-012-0254-3.

Dziegielewski, B. and Baumann, D.D. (2011) Predicting future demands for water, in *Treatise on water science*, vol. 1, edited by P.Wilderer, Rogers, P., Uhlenbrook, S., Frimmel, F., Hanaki, K., Vereijken, T., pp. 189-199, Elsevier.

Eisenhart, K.M. (1989) Building theories from case research, *The Academy of Management Review*, 14(4), 532-550.

Environment Agency, Ofwat, Defra and the Welsh Government (2012) Water resources planning guideline – the technical methods and instructions, EA, Bristol.

CH2MHill (2013), Water Resources Management Plans 2019 – Preparing for the Future (Main Report), report for the Environment Agency, CH2MHill, Swindon, UK.

Erlanger, P. and Neal, B. (2005), Framework for urban water resource planning, WSAA Occasional Paper No. 14 - ISBN 1 920760 19 9, Water Services Association of Australia (WSAA), Melbourne, Australia.

European Commission (1992) Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora, OJ L 206, p.7, EC, Brussels.

Fiering, M.B. (1976) The role of systems analysis in water program development, *Natural Resources Journal*, 16, 759-771.

Fiering, M.B. (1997) The real benefits from synthetic flows: reflections on 25 years of with the Harvard water program, in *Reflections on hydrology: science and practice*, edited by N.Buras, American Geophysical Union, Washington, D.C.

Feser, F., Rockel, B., von Storch, H., Winterfeldt, J. and Zahn, M. (2011), Regional climate models add value to global model data: A review and selected examples, *Bulletin of the American Meteorological Society*, 92.

Foddy, W. (1993) Constructing questions for interviews and questionnaires – theory and practice in social research, Cambridge University Press, Cambridge, UK.

Fowler, H., Blenkinsop, S. and Tebaldi, C. (2007), Linking climate change modeling to impacts studies: Recent advances in downscaling techniques for hydrological modeling, *International Journal of Climatology*, 27, 1547-1578.

Frick, D. M., Bode, D., and Salas, J. D. (1990) Effect of drought on urban water supplies. I: Drought analysis, *Journal of Hydraulic Engineering*, 116(6), 733-753, doi: 10.1061/(ASCE)0733-9429(1990)116:6(733).

Ghile, Y. B., Taner, M. Ü, Brown, C., Grijsen, J. G. and Talbi, A. (2014), Bottom-up climate risk assessment of infrastructure investment in the Niger River Basin, *Climatic Change*, 122(1-2), 97-110. doi: 10.1007/s10584-013-1008-9.

Giorgi, F. and Linoello, P. (2008), Climate change projections for the Mediterranean region, *Global and Planetary Change,* 63, 90-104.

Giuliani, M., J. D. Herman, A. Castelletti, and P. Reed (2014), Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management, *Water Resources Research*, 50, 3355–3377, doi:10.1002/2013WR014700.

Gober, P. (2013) Getting outside the water box: The need for new approaches to water planning and policy. *Water Resources Management*, 27(4), 955-957.

Groves, D. G., D. Yates, and C. Tebaldi (2008), Developing and applying uncertain global climate change projections for regional water management planning, *Water Resources Research*, 44, W12413, doi:10.1029/2008WR006964.

Gu, J. J., G. Guo and H. Huang (2013) Inexact stochastic dynamic programming method and application to water resources management in Shandong China under uncertainty, *Stochastic Environmental Research and Risk Assessment*, 27, 5, 1207-1219. doi: 10.1007/s00477-012-0657-y.

Hall, J.W. and Borgomeo, E. (2013) Risk-based principles for defining and managing water security, *Philosophical Transactions of the Royal Society A*, 371:20120407, doi: 10.1098/rsta.2012.0407.

Hall, J. W., Lempert, R. J., Keller, K., Hackbarth, A., Mijere, C. and McInerney, D. J. (2012b), Robust climate policies under uncertainty: A comparison of Robust Decision Making and Info-Gap methods, *Risk Analysis*, 32(10), 1657–1672. doi: 10.1111/j.1539-6924.2012.01802.x.

Hall, J.W., Watts, G., Keil, M., de Vial, L., Street, R., Conlan, K., O'Connell, P.E., Beven, K.J., Kilsby, C.G. (2012a) Towards risk-based water resources planning in England and Wales under a changing climate, *Water and Environment Journal*, 26, 118-129, doi: 10.1111/j.1747-6593.2011.00271.x.

Hallegatte, S., Shah, A., Lempert, R., Brown, C. and Gill, S. (2012), Investment decision making under deep uncertainty: Application to climate change, *Research Policy Working Paper 6193*, Office of the Chief Economist, Sustainable Development Network, The World Bank, Washington, DC.

Harris, C.N.P., Quinn, A.D., Bridgeman, J. (2013) Quantification of uncertainty sources in a probabilistic climate change assessment of future water shortages, *Climatic Change*, 121(2), 317-329, doi: 10.1007/s10584-013-0871-8.

Hashimoto, T., Stedinger, J.R., Loucks, D.P. (1982a), Reliability, resilience, and vulnerability criteria for water resource system performance evaluation, *Water Resources Research*, 18 (1), 14-20, doi: 10.1029/WR018i001p00014.

Hashimoto, T., D. P. Loucks, and J. R. Stedinger (1982b) Robustness of water resources systems, *Water Resources Research*, 18(1), 21–26, doi:10.1029/WR018i001p00021.

Hazen, A. (1914) Storage to be provided in impounding reservoirs for municipal water supply, *Transactions of the American Society of Civil Engineers*, 77, 1539.

Held, I.M. and Soden, B.J. (2006), Robust responses of the hydrological cycle to global warming, *Journal of Climate*, 19, 5686-5699.

Hirsch, R. M., J. L. Cohon, and C. S. ReVelle (1977), Gains from joint operation of multiple reservoir systems, *Water Resources Research*, 13(2), 239–245, doi:10.1029/WR013i002p00239.

Hirsch, R. M. (2011), A Perspective on Nonstationarity and Water Management. *Journal of the American Water Resources Association*, 47: 436–446. doi: 10.1111/j.1752-1688.2011.00539.x

Hurst, H. E. (1951) Long-term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers*, 116, pp. 770-808.

Jeuland, M., and D. Whittington (2014), Water resources planning under climate change: Assessing the robustness of real options for the Blue Nile, *Water Resources Research*, 50, 2086–2107, doi:10.1002/2013WR013705.

Kasprzyk, J. R., P. M. Reed, B. R. Kirsch, and G. W. Characklis (2009), Managing population and drought risks using many-objective water portfolio planning under uncertainty, *Water Resources Research*, 45, W12401, doi:10.1029/2009WR008121.

Kasprzyk, J. R., S. Nataraj, P. M. Reed, R. J. Lempert (2013) Many objective robust decision making for complex environmental systems undergoing change, *Environmental Modelling and Software*, 42, 55-71, doi:10.1016/j.envsoft.2012.12.007.

Kelley, P. & O'Brien, A. (eds) (2012), Source Scientific Reference Guide, eWater Cooperative Research Centre, Canberra, Australia.

Kilsby, C.G., Jones, P.D., Burton, A., Ford, A.C., Fowler, H.J., Harpam, C., James, P., Smith, A. and Wilby, R.L. (2007) A daily weather generator for use in climate change studies, *Environmental Modelling and Software*, 22 (12), 1705–1719.

Klemeš, V., R. Srikanthan, and T. A. McMahon (1981), Long-memory flow models in reservoir analysis: What is their practical value?, *Water Resources Research*, 17(3), 737–751, doi:10.1029/WR017i003p00737.

Klemeš, V. (1987) One hundred years of applied storage reservoir theory, *Water Resources Management*, 1(3), 159-175. doi: 10.1007/BF00429941.

Klemeš, V. (2001) Risk analysis: The unbearable cleverness of bluffing, in *Risk, reliability, uncertainty, and robustness of water resource systems*, chapter 3, edited by J. J. Borardi and Z. W. Kundzewicz, pp.101-110, Cambridge University Press, New York.

Korteling, B., Dessai, S., Kapelan, Z. (2013) Using information-gap decision theory for water resources planning under severe uncertainty, *Water Resources Management*, 27, 1149-1172, doi: 10.1007/s11269-012-0164-4.

Koutsoyiannis, D., Efstratiadis, A., Georgakakos, K. P. (2007) Uncertainty assessment of future hydroclimatic predictions: A Comparison of Probabilistic and Scenario-Based Approaches. *Journal of Hydrometeorology*, 8, 261–281. doi:10.1175/JHM576.1.

Koutsoyiannis, D., Montanari, A., Lins, H. F. and Cohn, T. A. (2009) Climate, hydrology and freshwater: towards an interactive incorporation of hydrological experience into climate research, *Hydrological Sciences Journal*, 54(2), 394-405, doi: 10.1623/hysj.54.2.394.

Koutsoyiannis, D. (2011), Hurst-Kolmogorov Dynamics and Uncertainty. *JAWRA Journal of the American Water Resources Association*, 47: 481–495. doi: 10.1111/j.1752-1688.2011.00543.x.

Kundzewicz , Z. W., Mata , L. J., Arnell , N.W., Döll , P., Jimenez , B., Miller, K., Oki , T., Şen, Z., Shiklomanov, I. (2008) The implications of projected climate change for freshwater resources and their management, *Hydrological Sciences Journal*, 53(1), 3-10, doi: 10.1623/hysj.53.1.3.

Kundzewicz, Z.W. and Stakhiv, E.Z. (2010) Are climate models ready for prime time in water resources management, or is more research needed?, *Hydrological Sciences Journal*, 55(7), 1085-1089.

Kundzewicz, Z. W. (2011), Nonstationarity in Water Resources – Central European Perspective. *Journal of the American Water Resources Association*, 47: 550–562. doi: 10.1111/j.1752-1688.2011.00549.x

Lambert, A. (1988) An introduction to operational control rules using the 10-component method. BHS Occasional paper No 1 (available for download from the British Hydrological Society).

Lane, W. L. (1979), Applied stochastic techniques (Last computer package); User manual, Division of Planning Technical Services, U.S. Bureau of Reclamation, Denver, Colorado.

Lempert, R. J. and Groves, D. G. (2010), Identifying and evaluating robust adaptive policy responses to climate change for water management agencies in the American west, *Technological Forecasting and Social Change* 77(6), 960-974. doi: 10.1016/j.techfore.2010.04.007.

Lempert, R.J., Popper, S.W. and Bankes, S.C. (2003) Shaping the next one hundred years: new methods for quantitative, long-term policy analysis, RAND, Santa Monica, CA.

Lind, R.C. (1997) Intertemporal equity, discounting, and economic efficiency in water policy evaluation. In: Climate change and water resources planning criteria, K.D. Frederick, D.C. Major, and E.Z. Stakhiv (Editors). Kluwer Publishers, Netherlands.

Lins and Cohn (2011) Stationarity: Wanted dead or alive, *Journal of the American Water Resources Association*, 46(3), 475-480.

Lins and Stakhiv (1998) Managing the nation's water in a changing climate, *Journal of the American Water Resources Association*, 34(6), 1255-1264, doi: 10.1111/j.1752-1688.1998.tb05429.x.

Lopez, A., Fung, F., New, M., Watts, G., Weston, A. and Wilby, R.I. (2009) From climate model ensembles to climate change impacts: A case study of water resources management in the south west of England, *Water Resources Research,* 45, W08419, doi: 10.1029/2008WR007499.

Loucks, D. (1992) water resource systems models: Their role in planning." *Journal of Water Resources Planning and Management*, 118(3), 214–223.

Loucks, D.P., van Beek, E., Stedinger, J.R., Dijkman, J.P.M. and Villars, M.T. (2005) Water resources systems planning and management – an introduction to methods, models and applications. Unesco publishing, Paris, France.

Lund, J.R. (2008) A risk analysis of risk analysis, *Journal of Contemporary Water Research and Education*, 40, 53-60. doi: 10.1111/j.1936-704X.2008.00028.x

Lund, J.R. (2012) Provoking more productive discussion of wicked problems, *Journal of Water Resources Planning and Management*, 138(3), 193-195.

Maass, A. and Hufscmidt, M.A. (1959) In search of new methods for river system planning, *Journal of the Boston Society of Civil Engineers*, 46 (2), 99-125.

Maass, A., Hufschmidt, M. M., Dorfman, R., Thomas, H. A., Marglin, S. A. and Fair, G. M. (1962), Design of water-resource systems: New techniques for relating economic objectives, engineering analysis, and governmental planning, Harvard University Press, Cambridge, Massachusetts.

Manning, L.J., Hall, J.W., Fowler, H.J., Kilsby, C.G. and Tebaldi, C. (2009) Using probabilistic climate change information from a multi-model ensemble for water resources assessment, *Water Resources Research*, 45, W11411, doi: 10.1029/2007WR006674.

Matalas, N. (2012), Comment on the announced death of stationarity, *Journal of Water Resources Planning and Management*, 138(4), 311–312. doi: 10.1061/(ASCE)WR.1943-5452.0000215.

Matalas, N. C., and M. B. Fiering (1977), Water resource systems planning, in *Climate, Climatic Change and Water Supply*, pp. 99–110, National Academy of Sciences, Washington, D. C.

Matrosov, E.S., Harou, J.J. and Loucks, D.P. (2011) A computationally efficient open-source water resource system simulator – Application to the London and the Thames Basin, *Environmental Modelling and Software*, 26 (12), 1599-1610, doi:10.1016/j.envsoft.2011.07.013.

Matrosov, E.S., Padula, S., Harou, J. (2013a) Selecting portfolios of water supply and demand management strategies under uncertainty – contrasting economic optimisation and robust decision making approaches, *Water Resources Management*, 27, 1123-1148, doi: 10.1007/s11269-012-0118-x.

Matrosov, E. S., Woods, A. M. and Harou, J. J. (2013b), Robust decision making and info-gap decision theory for water resource system planning, *Journal of Hydrology*, 494, 43-58. doi: 10.1016/j.jhydrol.2013.03.006.

Mays, L.W. (2011) Water Resources Engineering, 2nd Edition, Wiley, New York.

McMahon, T. A., Pegram, G. G. S., Vogel, R. M. and Peel, M. C. (2007) Review of Gould-Dincer reservoir storage-yield-reliability estimates, *Advances in Water Resources*, 30(9), 1873-1882. doi: 10.1016/j.advwatres.2007.02.004.

McMahon, T. A., Peel, M. C., and Karoly, D. J. (2014) Uncertainty in runoff based on Global Climate Model precipitation and temperature data – Part 1: Assessment of Global Climate Models, Hydrology and Earth System Sciences Discussions, 11, 4531-4578, doi:10.5194/hessd-11-4531-2014, 2014.

McMahon, T. A., Peel, M. C., Pegram, G. G. S. and Smith, I. N. (2011) A simple methodology for estimating mean and variability of annual runoff and reservoir yield under present and future climates, *Journal of Hydrometeorology*, 12(1), 135-146. doi: 10.1175/2010JHM1288.1.

Meehl, G. A., T. L. Delworth, M. Latif, B. McAveney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor (2007), The WCRP CMIP3 multimodel data-set: A new era in climate change research, *Bulletins of the American Meteorological Society*, 88, 1383–1394, doi:10.1175/BAMS-88-9-1383.

Means, E III.,Laugier, M., Daw, J., Kaatz, L. and Waage, M. (2010) Decision support planning methods: Incorporating climate change uncertainties into water planning, Water Utility Climate Alliance, San Francisco.

Memon, F.A. and Butler, D. (2006) Water consumption trends and demand forecasting techniques, in *Water demand management*, edited by Butler, D. and Memon, F.A., pp. 1-25, IWA Publishing, London.

Milly, P.C.D., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P. and Stouffer, R.J. (2008) Stationarity is dead: whither water management? *Science*, 319 (5863), 573-574, doi: 10.1126/science.1151915.

Moody, P., and C. Brown (2013), Robustness indicators for evaluation under climate change: Application to the upper Great Lakes, *Water Resources Research*, 49, 3576–3588, doi:10.1002/wrcr.20228.

Montaseri, M. and Adeloye, A. (2002) Effects of integrated planning on capacity-yield-performance functions." *Journal of Water Resouces Planning and Management*, 128(6), 456–461.

Mortazavi, M., Kuczera, G. and Cui, L. (2012) Multiobjective optimization of urban water resources: Moving toward more practical solutions, *Water Resources Resources*, 48, W03514, doi:10.1029/2011WR010866.

Nakićenović, N., Alcamo, J., Davis, G., de Vries, H. J M., Fenhann, J., Gaffin, S., Gregory, K., Grubler, A., Jung, T. Y., Kram, T., La Rovere, E. L., Michaelis, L., Mori, S., Morita, T., Papper, W., Pitcher, H., Price, L., Riahi, K., Roehrl, A., Rogner, H. H., Sankovski, A., Schlesinger, M., Shulka, P., Smith, S., Swart, R., van Rooijen, S., Victor, N. and Dadi, Z. (2000), Emissions scenarios: A special report of Working Group III of the International Panel on Climate Change, Cambridge University Press, Cambridge, UK.

New, M., Lopez, A., Dessai, S. and Wilby, R. (2007) Challenges in using probabilistic climate change information for impacts assessments: An example from the water sector, *Philosophical Transactions of the Royal Society*, 365 (1857), 2117-2131, doi: 10.1098/rsta.2007.2080.

New Jersey Department of Environmental Protection (2011) Estimating the safe yield of surface water supply reservoirs systems, Guidance Manual, New Jersey Department of Environmental Protection, Trenton, New Jersey.

Nilsen, V., Lier, J. A., Bjerkholt, J. T. and Lindholm, O. G. (2011), Analysing urban floods and combined sewer overflows in a changing climate, *Journal of Water and Climate Change*, 2, 260-271.

Oxford Scientific Software (2008), A guide to Aquator, Oxford Scientific Software, Oxford.

Pallottino, S., Sechi, G.M. and Zuddas, P. (2005) A DSS for water resources management under uncertainty by scenario analysis, *Environmental Modelling and Software*, 20, 1031-1042.

Perraud, J. M., Podger, G. M., Rahman, J. M. and Vertessy (2003) A new rainfall runoff software library, *Proceedings of MODSIM*, vol. 4.

Prudhomme, C., Dadson, S., Morris, D., Williamson, J., Goodsell, G., Crooks, S., Boelee, L., Davies, H., Buys, G., Lafon, T. and Watts, G. (2012) Future flows climate: an ensemble of 1-km climate change projections for hydrological application in Great Britain, *Earth System Science Data Discussions*, 5, 475-490, doi: 10.5194/essdd-5-475-2012.

R Core Team (2014), R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

Ratnayaka, D. D., Brandt, M. J. and Johnson, K. M. (2009) Twort's water supply, Sixth edition, Elsevier, Oxford.

Reuss, M. (2003), Is it time to resurrect the Harvard Water Program? *Journal of Water Resources Planning and Management*, 129 (5), 357-360, doi: 10.1061/(ASCE)0733-9496(2003)129:5(357).

Ricketts, J. H. and Page, C. M. (2007) A web based version of OzClim for exploring climate change impacts and risks in the Australian regions, CSIRO Marine and Atmospheric Research, Aspendale, Victoria.

Rippl, W. (1883) The capacity of storage reservoirs for water supply, In *Proceedings of the Institute of Civil Engineers*, 71, 270-278.

Rogers, P.P., and M.B. Fiering (1986), Use of systems analysis in water management, *Water Resources Research*, 22(9S), 146S–158S, doi: 10.1029/WR022i09Sp0146S.

Rogers, P.P. (1997) Engineering design and uncertainties related to climate change, *Climate Change*, 37, 229-242.

Rosenberg, D. (2012) Near-optimal water management to improve multi-objective decision making, *International Conference of Environmental Modelling and Software, Managing Resources of a Limited Planet*, 6th Biennial meeting, Leipzig, Germany.

Rosenberg, D. and Madani, K. (2014). Water resources systems analysis: A bright past and a challenging but promising future. *Journal of Water Resources Planning and Management*, 140(4), 407-409.

Rush, P.V., Murphy, T., Mayer, R.A., Johnstone, T., Olivio, D. and Woods, H.J. (2011) Water system safe yield calculation, New York City Department of Environmental Protection, New York.

Salas, J. D. (2013), Discussion: “pragmatic approaches for water management under climate change uncertainty” by Eugene Stakhiv, *Journal of the American Water Resources Association*, 49(2), 475-478. doi: 10.1111/jawr.12026.

Salas, J., Rajagopalan, B., Saito, L., and Brown, C. (2012), Special section on climate change and water resources: climate nonstationarity and water resources management., *Journal of Water Resources Planning and Management*, 138(5), 385–388. doi: 10.1061/(ASCE)WR.1943-5452.0000279.

Southern Water (2013), Revised draft Water Resources Management Plan 2015 – 2040: Technical report, Brighton, UK.

Srikanthan R. (2005), Stochastic Generation of daily rainfall at a number of sites, Cooperative Research Centre for Catchment Hydrology, Technical report 05/7, Canberra, Australia.

Smithers, H. and Walker, S. (1997) Reassessment of water resources in northwest England as a result of the 1995-1996 drought, in *Sustainability of water resources under increasing uncertainty*, edited by Robbjerg, D., Boutayeb, N., Gustard, A., Kundzewicz, Z.W. and Rasmussen, P.F., pp. 173-181, IAHS, Wallingford, Oxfordshire.

Stainforth, D. A., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D. J., Kettleborough, J. A., Knight, S., Martin, A., Murphy, J. M., Piani, C., Sexton, D., Smith, L. A., Spicer, R. A., Thorpe, A. J. and Allen, M. R. (2005), Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature*, 433, 403-406. doi: 10.1038/nature03301.

Stainforth D. A., Allen, M. R., Tredger, E. R. and Smith, L. A. (2007), Confidence, uncertainty and decision-support relevance in climate predictions, *Philosophical Transactions of the Royal Society A*, 365, 2145-2161. doi: 10.1098/rsta.2007.2074.

Stakhiv, E. Z. (2011), Pragmatic approaches for water management under climate change uncertainty, *Journal of the American Water Resources Association*, 47(6), 1183-1196. doi: 10.1111/j.1752-1688.2011.00589.x.

State of Queensland Department of Environmental and Resource Management (2010) Far north Queensland—Regional water supply strategy, Report #28725, State of Queensland Department of Environmental and Resources Management, Brisbane.

Steinschneider, S., and C. Brown (2013), A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, *Water Resources Research*, 49, 7205–7220, doi:10.1002/wrcr.20528.

Sveinsson, O. G. B., Salas, J. D., Boes, D. C. and Pielke, R. A. (2003), Modeling the dynamics of long term variability of hydroclimatic processes, *Journal of Hydrometeorology*, 4(3), 489-505. doi: 10.1175/1525-7541(2003)004.

Sveinsson, O. G. B., Salas, J. D., Lane, W. L. and Frevert, D. K. (2007), Stochastic Analysis, Modeling, and Simulation (SAMS) Version 2007; User's manual, *Technical report no. 11*, Colorado State University, Fort Collins, Colorado.

Sudler, C. E. (1927) Storage required for regulation of stream flow, *Transactions of the American Society of Civil Engineers*, 91, 622.

Taleb, N.N. (2007) The black swan: the impact of the highly improbable, Random House, New York.

Tennessee Department of Environment and Conservation (2013) Regional water resources planning guidelines for Tennessee, TdEC, Nashville, Tennessee.

Thames Water (2013) Revised Draft Water Resources Management Plan 2015 – 2040, Reading, UK.

Uff, J. (1996) Water supply in Yorkshire, Report of the Independent Commission of Inquiry, Keating Chambers, London.

Umgeni Water (2013) Infrastructure Master Plan 2013: 2013/2014 – 2043/44 Volume 1, KwaZulu-Natal, Pietermaritzburg, South Africa.

UK Infrastructure Transition Research Consortium (2014) National Infrastructure Assessment: Analysis of options for infrastructure provision in Great Britain—Interim results, January 2014, Oxford, UK.

UK Water Industry Research (2000) A unified methodology for the determination of deployable output from Water Sources (UKWIR WR-18), UKWIR, London, UK.

UK Water Industry Research (2012a) *Water resources planning tools 2012; Evaluation of new methods for 2019 and beyond*, UKWIR, London, UK.

UK Water Industry Research (2012b) *Water resources planning tools 2012; Deployable Output (DO) report*, UKWIR, London, UK.

United Utilities PLC (2013) draft Water Resources Management Plan—Technical Report: Water source yields and supply capability, United Utilities PLC, Warrington, UK.

Vamvakeridou-Lyroudia, L.S., Morley, M.S., Bicik, J., Green, C., Smith, M.A. and Savic, D.A. (2009) Aquator GA: Integrated optimization for reservoir operation using multiobjective genetic algorithms, *Proceedings of the 10th International Conference of Computing and Control for the Water Industry*, Sheffield, UK, pp 493-500.

van Dijk, A. I. J. M., H. E. Beck, R. S. Crosbie, R. A. M. de Jeu, Y. Y. Liu, G. M. Podger, B. Timbal, and N. R. Viney (2013), The Millennium Drought in southeast Australia (2001–2009): Natural and human causes and implications for water resources, ecosystems, economy, and society, *Water Resources Research*, 49, doi:10.1002/wrcr.20123.

Varlet, M. (1923), "Etude Graphique des Conditions d'Exploitation d'un Reservoir de Regularisation," Annales des Ponts et Chaussees, Partie Technique, 93e Annee, 1 le erie-Time 64, Fasc. 4. Tome 2, July-August.

Vidal, J.-P., and S. Wade (2008), A framework for developing high-resolution multi-model climate projections: 21st century scenarios for the UK, *International Journal of Climatology*, 28, 843–858, doi:10.1002/joc.1593

Vogel, R. M., and Stedinger, J. R. (1987), Generalized storage-reliability-yield relationships, *Journal of Hydrology*, 89(3), 303-327.

Vogel, R. M., and R. A. Bolognese (1995), Storage-reliability-resilience-yield relations for over-year water supply systems, *Water Resources Research*, 31(3), 645–654, doi:10.1029/94WR02972.

Vogel, R. M., Bell, C. J., Suresh, R. R. and Fennessey, N. M. (2001) Regional assessment of the impacts of climate change on the yield of water supply systems, in *Risk, reliability, uncertainty, and robustness of water resource systems*, chapter 12, edited by J. J. Borardi and Z. W. Kundzewicz, pp.101-110, Cambridge University Press, New York.

Voss, C., Tsekriktsis, N. and Frohlich, M. (2002) Case research in operations management, *International Journal of Operations and Production Management*, 22(2), 195-219.

Waage, M. D. and Kaatz, L. (2011), Nonstationary water planning: An overview of several promising planning methods, *Journal of the American Water Resources Association*, 47(3), 535-540. doi: 10.1111/j.1752-1688.2011.00547.x.

Walker, S. and Smithers, H. (1998) A review of the 1995-96 drought in the North West, *Water and Environment Journal*, 12 (4), 273-279, doi: 10.1111/j.1747-6593.1998.tb00185.x.

Water Research Foundation (2013) Defining and enhancing the safe yield of a multi-use, multi-reservoir water supply, Tailored Collaboration – Web Report #4304, Water Research Foundation, Denver, Colorado.

Watts, G., von Christierson, B., Hannaford, J., Lonsdale, K. (2012) Testing the resilience of water supply systems to long droughts, *Journal of Hydrology*, 414-415, pp. 255-267. doi: 10.1016/j.jhydrol.2011.10.038.

Weaver, C. P., Lempert, R. J., Brown, C., Hall, J. A., Revell, D. and Sarewitz, D. (2013), Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks, *WIREs Climate Change*, 4(1), 39–60. doi: 10.1002/wcc.202.

Westcott, R. (2004) A scenario approach to demand forecasting, *Water Science and Technology: Water Supply*, 4 (3), 45-55.

Whetton, P., Hennessy, K., Clarke, J., McInnes, K. and Kent, D. (2012), Use of representative climate futures in impact and adaptation assessment, *Climatic Change*, 115, 433-442.

Wilby, R. L. (2010) Evaluating climate model outputs for hydrological applications – Opinion, *Hydrological Sciences Journal*, 55(7), 1090–1093.

Yin, R.K. (1994) Case study research: design and methods, Second Edition, Sage, Thousand Oaks, California.

Zarghami, M. and H. Hajykazemian (2013) Urban water resources planning by using a modified particle swarm optimization algorithm, *Resources, Conservation and Recycling*, 70, 1-8. doi: 10.1016/j.resconrec.2012.11.003.